

**READABILITY EASE OF
ONLINE HEARING RELATED INFORMATION IN HINDI**

Seema Diwan

**A thesis submitted in partial fulfilment of the
requirements for the Degree of
Master of Audiology
in the Department of Communication Disorders
at the University of Canterbury
2017**

ACKNOWLEDGEMENTS

After an intense journey of learning this year, now is the time to write a note of thanks. This period not only improved my scientific expertise, it also helped me learn at the personal level. Continuous writing for months has improved my writing skills, and my patience towards other aspects of life. I would like to acknowledge those people who supported me during this journey.

I would first like to thank my supervisor, Dr Rebecca Kelly-Campbell who supported me through the journey of this programme and helped me throughout with my research and writing. I would like to thank her for her kindness, patience and determination in always driving me, which kept me on the right track the whole year. She has immense knowledge of research and is an academic genius. I am really grateful to her for being my thesis mentor.

Besides my supervisor, I would like to thank the rest of my supervisors in the Department of Communication Disorders: Dr Greg O'Beirne, Jonathan Grady and Nicole Borland for being continuous supportive and providing the necessary resources for the research.

My sincere thanks to my classmates for their encouragement and sympathetic ear. We were able to discuss our findings and look out for each other during bad times. Finally, there is my spouse, Dr S. Yeh who stood with me all the way through this path and provided me with the strength I needed to accomplish this huge goal.

“Kindness is a language that the deaf can hear and the blind can see”

(Mark Twain)

ABSTRACT

Purpose: The purpose of the study was to assess the readability of hearing-related internet information in the Hindi language.

Methods: Five keywords which were commonly used during the Internet search related to hearing problems and hearing impairment were identified by native Hindi speakers. These key terms were entered one by one into Google Bharat, the Hindi version of Google India. The uniform resource locators (URLs) were recorded for the first ten webpage results obtained after entering each search phrase. Each webpage was assessed according to inclusion and exclusion criteria. Finally, 25 webpages were analysed by using a readability computational tool. RH1 and RH2 formulas (Sinha, Sharma, Dasgupta, & Basu, 2012) were used to calculate the Readability Grade Level (RGL). The paragraphs with lowest and highest readability level were identified and used for a cloze test. Ten participants were recruited after applying inclusion and exclusion criteria. They were instructed to complete the cloze test.

Results: The mean RGL of hearing-related webpages published in Hindi was not significantly different from the recommended value. A significant difference between the mean RGL for webpages of different origins was identified. The mean RGL calculated by RH1 was significantly higher than the mean RGL calculated by RH2; but there was a significant and positive correlation between the RGL values calculated by RH1 and RH2. No significant differences in cloze scores were found between the paragraph with the high RGL and the paragraph with the low RGL.

Conclusion: The RGL calculated by the formulas was within the recommended value, which means the hearing-related material available on the Internet in Hindi is easy to read. However, the results of readability ease calculated by the cloze test suggested that the paragraphs with

maximum RGL and minimum RGL were not significantly different from each other in the level of understanding.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF TABLES	viii
LIST OF FIGURES	ix
LIST OF ABBREVIATIONS	x
1 CHAPTER ONE: INTRODUCTION.....	1
1.1 Study overview	1
1.2 Hearing impairment.....	2
1.2.1 Overview	2
1.2.2 Prevalence of hearing impairment.....	4
1.2.3 Impact of hearing impairment	6
1.3 Management of Hearing Impairment	6
1.3.1 Effects of management of hearing impairment.....	6
1.3.2 Factors influencing success of rehabilitation of hearing impairment	8
1.4 Online health information.....	10
1.4.1 Access to online health information and its use globally	10
1.4.2 Access to online health information and its use in India	12
1.4.3 Access to online health information in Hindi speaking population	13
1.5 Health literacy.....	14
1.5.1 Health literacy globally	14
1.5.2 Health literacy in India	15
1.6 Quality of online health information.....	16
1.6.1 Readability	16
1.6.2 Other measures for quality control	18
1.6.3 Cloze test	20
1.7 Readability analyses.....	22
1.7.1 Readability analysis in Hindi.....	22
1.7.2 Features of Hindi language.....	23
1.7.3 Differences between Hindi and English	23
1.7.4 Readability computational measures available in Hindi.....	24
1.8 Study rationale.....	29
1.9 Aims and hypothesis	29
2 CHAPTER TWO: METHODS.....	31
2.1 Part One (Readability).....	31
2.1.1 Identification of search terms	31
2.1.2 Internet search	32
2.1.3 Inclusion and exclusion criteria.....	32
2.1.4 Webpage information	32
2.1.5 Selection of a paragraph	33
2.1.6 Readability computational tool.....	33

2.2	Part Two (Cloze test).....	34
2.2.1	Ethics Approval.....	34
2.2.2	Selection of paragraphs	34
2.2.3	Participants	34
2.2.4	Procedures	35
2.2.5	Measures.....	35
2.2.6	Planned statistical analysis	36
3	CHAPTER THREE: RESULTS	38
3.1	Overview	38
3.2	Part 1 (Readability).....	38
3.2.1	Hypothesis 1	39
3.2.2	Hypothesis 2	40
3.2.3	Hypothesis 3	41
3.2.4	Hypothesis 4.....	41
3.2.5	Hypothesis 5	41
3.3	Part 2 (Cloze test).....	42
3.3.1	Participants	42
3.3.2	Hypothesis 6	43
4	CHAPTER FOUR: DISCUSSION	44
4.1	Introduction	44
4.2	Relation to Literature and clinical significance	44
4.2.1	Readability of Online Hearing-related Health Information in Hindi	44
4.2.2	Testing readability ease by a cloze test.....	47
4.3	Clinical implications.....	49
4.4	Study limitations	50
4.4.1	Readability search and webpages in Hindi.....	50
4.4.2	Readability Formulas	51
4.4.3	Cloze test.....	52
4.4.4	Other limitations.....	53
4.5	Future research	53
4.5.1	New models for readability formulas	53
4.5.2	Health On the Net (HON) Code and DISCERN.....	53
4.5.3	Focus on the readability of important sections of the webpage.....	54
4.5.4	Focus on readability and comprehension	54
4.6	Conclusions	55
5	REFERENCE LIST.....	56
APPENDIX A.....	71	
APPENDIX A.1 - ETHICS APPROVAL LETTER.....	71	
APPENDIX A.2. DEMOGRAPHIC QUESTIONNAIRE	72	
APPENDIX A.3. INFORMATION SHEET.....	74	
APPENDIX A.4. CONSENT FORM.....	76	
APPENDIX A.5. CLOZE PROCEDURE	78	

LIST OF TABLES

Table 1: Internet World Statistics by World Internet Usage and Population Statistics.....	11
Table 2: Cloze and multiple-choice scores comparison.....	21
Table 3: Inclusion criteria and exclusion criteria for participants.....	34
Table 4: Planned statistical analyses of each hypothesis	37
Table 5: Means, standard deviations, minimum, maximum values, and sample sizes for each readability formula	39
Table 6: Kruskal–Wallis Test Results	40
Table 7: Descriptive summary of participants	42
Table 8: Cloze score data obtained from participants and their descriptive summary	43

LIST OF FIGURES

Figure 1: Graphical User Interface (GUI) of the computational tool	26
Figure 2: Box Plot of Reading Grade Level of Hindi webpages. RH1= Readability Hindi1. RH2= Readability Hindi 2.....	39
Figure 3: Relationship between RH1 and RH2 Formulas. Blue dots represent reading grade level (RGL) of each webpage and linear regression line is the red dotted line.	41

LIST OF ABBREVIATIONS

ASHA	American Speech-Language-Hearing Association
ASL	Average Sentence Length
AWL	Average Word Length
ccTLD	Country Coded Top Level Domain
CRIE	Chinese Readability Index Explorer
dB HL	Decibel Hearing Level
F-K	Flesch-Kincaid
FRE	Flesch Reading Ease
HON	Health On the Net
IBM	International Business Machine
JUK	Jukta-akshars
kHz	Kilohertz
PC	Personal Computer
PSW	Polysyllabic Words
PSW30	Polysyllabic words per 30 sentences
RGL	Reading Grade Level
RH1	Readability Hindi 1
RH2	Readability Hindi 2
SMOG	Simple Measure of Gobbledygook
SPSS	Statistical Package for the Social Sciences
URL	Uniform Resource Locator
WHO	World Health Organisation

1.1 Study overview

Healthcare information available online is becoming an important source of information for consumers with health conditions such as diabetes, hypertension and hearing impairment (Cline & Haynes, 2001). This information is used by consumers to decide what further actions to take, and how to take good care of themselves in terms of their health (Fiksdal et al., 2014). According to the data provided by International Telecommunication Union (2016), Internet users are increasing globally, within which China and India are becoming a huge market. An interesting fact about the Indian population observed by Bhattacharya (2017) is that they trust the content available in Hindi more than in English, which clearly makes it important to look at the readability ease of the online health-related information available in Hindi.

The majority of research has shown that the readability level of online health related material is higher than the recommended readability level by experts which is fifth or sixth grade level (Walsh & Volsko, 2008). The same tendency has been shown by the research available in audiology, which means the readability of hearing-related information online is poor (Atcherson et al., 2014). To date, there is no study available in the literature identifying the readability ease of online hearing-related information in Hindi. The present study aims to identify the readability level of online hearing-related material in Hindi by using readability formulas and a cloze test.

In the next subsection, the literature review will provide the outline of hearing impairment, advantages of rehabilitation, health literacy and Internet usage in Hindi, and readability formulas in Hindi, before defining the aims and hypothesis of the study.

1.2 Hearing impairment

1.2.1 Overview

Hearing impairment is defined by the World Health Organization (2001) as the complete or partial inability to hear, which can be due to structural or functional abnormalities of the auditory system. A significant impact of hearing impairment has been observed in adults by Finitzo and Crumley (1999). They found that adults with hearing impairment face a great challenge while communicating with people around them, mainly due to lack of consideration by others. This can further lead to social isolation, depression and functional disability (Finitzo & Crumley, 1999).

The human hearing organ or auditory system converts mechanical signals from the surroundings into electrical signals with the help of various parts of the peripheral hearing organ such as the outer ear, middle ear, and inner ear, and the central auditory system such as the neural pathways, auditory brain stem and auditory cortex (Laforge, Spector, & Sternberg, 1992). The peripheral hearing system comprises the outer ear, middle ear and inner ear. When sound reaches the outer ear or pinna, it is directed towards the tympanic membrane which starts vibrating (Newton & Shah, 2013). The vibrations of the tympanic membrane cause the middle ear ossicles (malleus, incus and stapes) to vibrate, and energy is transferred to the inner ear (Mulrow et al., 1990). This energy causes movement of the cochlear fluid and basilar membrane, which stimulates a number of nerve endings, causing nerve impulses to be transmitted to the auditory cortex through various nuclei along the auditory neural pathway. The neural impulses are interpreted by the brain and a symbol, word, picture, or recognised signal is created (Møller, 2012).

Hearing impairment can be caused by any disruption in the above-mentioned structures; disruptions can be anatomical, pathophysiological or environmental, such as trauma, noise

exposure, and infections (Møller, 2012). The cause of hearing impairment can vary between the adult and paediatric population. Previous research has established that the cause of hearing impairment occurring in the paediatric population is 50% environmental (caused by infections, pregnancy complications, ototoxic antibiotics) and 50% hereditary (Musiek & Baran, 1986). However, the cause of hearing impairment in adults is more commonly due to ageing, resulting in pathophysiological alterations in the hearing organ, such as reduced function of the cochlear hair cells and diminished sound processing capability in neural pathways (Musiek & Baran, 1986).

Three types of hearing impairments have been classified by researchers, based on the location of the cause or problem associated with the hearing organ: conductive, sensorineural and mixed. Conductive hearing impairment refers to a disturbance in the conduction of the sound waves occurring in the outer ear or the middle ear. Cerumen, ossicular chain deformities and their fixation, and middle ear infection can cause conductive hearing impairment (Steiger, 2015). This type of hearing impairment is often known as temporary hearing impairment because it can typically be treated by different types of surgery and medication (Bolz, 2016). On the other hand, sensorineural hearing impairment is a permanent impairment because it cannot be treated by medication or surgery (Gordon-Salant, Frisina, Fay, & Popper, 2010). Sensorineural hearing impairment occurs because of inner ear dysfunction that hampers the conversion of sound wave vibrations to electrical signals (Martin & Clark, 2012). Mixed type of hearing impairment is a combination of conductive and sensorineural impairment in which middle ear and inner ear dysfunction occur. Audiologists describe a hearing impairment by type: (a) the affected part of the mechanism; (b) the degree of severity – intensity level and range of sounds which are not heard; and (c) configuration of hearing impairment – the frequencies at which hearing impairment has occurred (American Speech-Language-Hearing Association, 2015).

Hearing impairment is assessed by standard pure tone audiometry, which is the most common method used in hearing clinics for the assessment of hearing impairment in adults (Martin & Clark, 2012). In this assessment, the client responds to various pure tones at different intensity levels. The audiologist measures the softest level of each pure tone to which the client responds. Play audiometry or conditioned play audiometry (Madell, 2014a) is a method used to assess the hearing of children between the ages of 2 and 5 years. This technique utilises behavioural conditioning. Visual reinforcement audiometry (Madell, 2014b) is used to assess the hearing of children who are in the cognitive age range from 6 to 36 months. This technique practices a conditioned head-turning response that is modelled by the examiner's control of a stimulus-reinforcement model.

1.2.2 Prevalence of hearing impairment

According to recent data provided by the World Health Organization (2017), 360 million people, (5% of the whole global population), have disabling hearing impairment (greater than 40 dB HL in the better hearing ear). Out of these 360 million people with hearing impairment, 328 million are adults and 32 million are children (Schlauch & Nelson, 2015). Stevens and colleagues reviewed 42 studies from 29 countries, published from 1973 to 2010, to estimate the global prevalence of hearing impairment (Stevens et al., 2013). The results showed that approximately 1.4% of children aged 5-14 years old are affected globally by hearing impairment. This percentage increased with an increase in age, and hearing impairment was found to be the most prevalent in South Asia where the rate was 17% (in adults older than 15 years).

The incidence and prevalence of hearing impairment is high in India (Garg, Chadha, Malhotra, & Agarwal, 2009). According to the Rehabilitation Council of India Act, 1992, "hearing handicap" is defined as "hearing impairment of 70 dB and above, in better ear or total loss of hearing in both ears". In 1995, The Persons with Disabilities (PWD) Act included a

wider range of degrees of hearing impairment. However, it still did not include people with a moderate degree of hearing impairment. According to a document released by the Rehab Council of India (1995), a person can be defined as disabled “who has a minimum of 60dB HL of hearing impairment in the better ear in speech conversation frequencies” (p. 100). Census 2001 (Office of the Registrar General & Census Commissioner, 2001) revealed a prevalence of 5.8% hearing disability in India. It is important to note this percentage did not include people with mild or moderate hearing impairment. In a paper published by American Speech-Language-Hearing Association ASHA (1981), hearing disability means “the determination of a financial award for the actual or presumed loss of ability to perform activities of daily living” (Summary, para. 1) and hearing “handicap” is the detriment due to hearing impairment on an individual’s daily life activities. In this report (American Speech-Language-Hearing Association ASHA, 1981), a person having average thresholds at 1, 2, 3 and 4kHz in the range of 25–75dBHL was considered to be hearing handicapped. It is clear that in India, the government does not consider people with mild to moderate hearing impairment as hearing handicapped.

The National Sample Survey Organization (2003) conducted a survey among Indian households and observed that hearing disability was the second most common disability in India. In rural communities, the disability was 10% of the population while it was 9% in urban communities. In addition, it was anticipated that out of 100,000 people, 291 people have hearing impairment. It was also observed that about 32% of the individuals with hearing impairment had profound hearing impairment, while 39% had severe hearing impairment. The survey also demonstrated that approximately 7% of children were born with hearing impairment. In cities, 62% of all the people in the survey reported that their hearing impairment was diagnosed at or above 60 years of age while this number was 56% in the villages.

The most common cause of hearing impairment in India found by a WHO survey (Varshney, 2016) was cerumen impaction (15.9%). Ageing was the next common aetiology found in this survey (10.3%). Middle ear infections were the cause of hearing impairment in 5.2% of the population. In half of the population with congenital hearing impairment, the causes of hearing impairment were environmental, such as viral infections, neonatal hypoxia, hyperbilirubinemia and meningitis. The cause of hearing impairment in the other half of the population was genetic (India, 1992).

1.2.3 Impact of hearing impairment

Hearing impairment can have a major impact on communication in daily life. The literature shows that hearing impairment has a negative impact on quality of life (Samuel & Babu, 2017). Dalton and colleagues (2003) revealed that severity of hearing impairment lowers the life quality in older adults. It can also result in social isolation, dependence and frustration (Varshney, 2016). In children, hearing impairment has been found to be related to delayed speech and language development (Garg et al., 2009). It can further impact their performance in school and psychological development. In some studies, it has been reported that parents of children with hearing impairment have higher anxiety levels and stress (World Health Organization, 2002). Hearing impairment can impact communication with family, friends and colleagues (Vishwambhar, 2015). The person with hearing impairment will respond slowly and will not be able to take part in active discussions. This will further impact their social life, and they will become more isolated (Vishwambhar, 2015).

1.3 Management of Hearing Impairment

1.3.1 Effects of management of hearing impairment

As the impact of hearing impairment brings negative social and psychological outcomes, management becomes an important consideration. Certain types of hearing

impairment can be managed medically and surgically. Examples include impairments caused by chronic infections (Lewis-Cullinan & Janken, 1990), sclerosed middle-ear ossicles (Williams, Chalmers, Stange, Chalmers, & Bowlin, 1993), atresia of the ear canal (Moss, Lin, & Cueva, 2016) and impacted cerumen (Yueh, Shapiro, MacLean, & Shekelle, 2003), sudden hearing impairment (Stokroos, Albers, & Tenvergert, 1998), and ototoxic drugs (Begg, Barclay, & Kirkpatrick, 2001). Management of hearing impairment that cannot be treated by surgical or medical procedures requires a team-based approach for care and rehabilitation (Pacala & Yueh, 2012).

A considerable quantity of literature has been published on the outcome of the provision of hearing aids (Mendel, 2007). For instance, a systematic review (Chisolm et al., 2007) of 16 studies demonstrated that regular use of hearing aids was related to emotional, social and psychological improvement in adults with sensorineural hearing impairment. Similarly, the effects of hearing aids on quality of life was examined for a sample of 1192 adults aged over 70 years (Appollonio, Carabellese, Frattola, & Trabucchi, 1996). The researchers established that uncorrected sensory deprivation was associated with a low quality of life and this kind of association was absent in people with corrected sensory deficit. Another similar type of survey was performed to assess the benefits of hearing aids on the quality of life (Kochkin & Rogin, 2000). In this survey, 2069 adults with hearing impairment completed a questionnaire that was developed from previously published and validated questionnaires. A separate questionnaire for family was also designed. The researchers concluded that hearing aids significantly help people with mild to severe hearing impairment by improving their physical and psychological health (Kochkin & Rogin, 2000). It is evident that hearing aids play a significant role in elevating the levels of satisfaction and life quality of people with hearing impairment (Cox & Alexander, 2002; Takahashi et al., 2007).

Other than hearing aids, the process of management of hearing impairment or aural rehabilitation comprises three other components: instruction (use of technology and listening environment), perceptual training (to improve communication) and counselling (to enhance participation) (Boothroyd, 2007). A prospective study was performed by Santos, Marangoni, de Andrade, Prestes, and Gil (2014) to identify long-term implications of an auditory training programme for adults with bilateral high frequency hearing impairment on various aspects of auditory function such as behavioural, subjective and electrophysiological function. The results of this study confirmed a significant improvement in communication in noisy environments, better skills for verbal speech sounds and temporal resolution. Evoked auditory potentials illustrated the presence of long latency components which were absent before the training programme. In another study conducted by Marques, Kozlowski, and Marques (2004), it was observed that an auditory rehabilitation programme including counselling, training and instruction associated with hearing aid fitting has a useful impact on an adult and his family dealing with hearing impairment.

1.3.2 Factors influencing success of rehabilitation of hearing impairment

Even though the advantages of hearing aids and other rehabilitation options are evident and established, most of the population with hearing impairment does not possess hearing aids. Numerous research studies have illustrated that only 14.6% to 20% of adults having objective or subjective hearing impairment use a hearing aid (Popelka et al., 1998; Stephens, Lewis, Davis, Gianopoulos, & Vetter, 2001). The existing rehabilitation choices and their use with satisfaction are guided by various factors.

In a systematic review carried out by Meyer and Hickson (2012), 22 articles examining factors influencing hearing aid adoption in older adults were reviewed. The authors concluded that people who are older, and have a moderate to severe hearing impairment, and identify their

family member or a friend as a supporter for the rehabilitation programme, are more likely to accept hearing aids.

Vestergaard Knudsen, Öberg, Nielsen, Naylor, and Kramer (2010) performed a review of available literature to evaluate factors affecting help-seeking, satisfaction, use and uptake of hearing aids. In this review, studies published between January 1980 and January 2009 were assessed. The patient journey was divided into three stages: pre-fitting, fitting and post-fitting. Of those, 22 studies focused on the factors affecting the pre-fitting stage, 2 studies focused on the fitting stage and 17 studies dealt with the factors playing a significant role in the post-fitting stage (up to 1 year). After selecting these studies, 31 factors were examined and discussed. These factors were demographic (age, gender), personal (expectation, motivating source, hearing sensitivity, hearing impairment), and external (counselling, cost). The results showed that for majority of the factors there was no consistency. However, gender, age and self-reported activity limitation illustrated consistent impact on the outcomes (satisfaction, use and uptake of hearing aids). An intriguing result from this study was that only one factor influenced all four outcome variables and that was self-reported hearing disability. This was an important finding and shows that self-reported auditory difficulty is an essential factor in successful aural rehabilitation.

In addition, existing research recognises the critical role played by the participation of the adults who are new to hearing aids, in a rehabilitation programme including communication strategies and their use (Hickson, Worrall, & Scarinci, 2007). It has also been observed that most people face great difficulty after finishing a hearing aid fitting and do not know what to do to solve these problems (Popelka et al., 1998). However, the regular use of hearing aids develops a positive attitude towards them and therefore increases satisfaction with the outcome (Eriksson-Mangold & Carlsson, 1991).

The rehabilitation programmes regarding hearing aid use and communication strategies are typically targeted towards a selected population with severe hearing impairment or who present a particular situation at work (Montano, Montano, & Spitzer, 2009). It indicates that not everybody who is fitted with hearing aids typically has access to these programmes (Öberg, Wänström, Hjertman, Lunner, & Andersson, 2009). To solve this issue, researchers have suggested using the Internet as a social medium to reach a larger number of people (Carlbring & Andersson, 2006). That study has established that the Internet can be regarded as a helpful counselling instrument for rehabilitation. The authors also suggested that through the Internet, material can be offered on a regular basis in a short time and to a distant location (Carlbring & Andersson, 2006). Other researchers have suggested that the Internet can provide cost-effective and modern methods to help people in their aural rehabilitation (Laplane-Lévesque, Kathleen Pichora-Fuller, & Gagné, 2006).

1.4 Online health information

The literature has provided evidence to suggest that consumers use the Internet as a source of health information. A study conducted by Diaz et al. (2002) to determine the use of the Internet for medical information by patients found that the majority of the patients routinely seek help from the Internet. In an another study, Tonsaker, Bartlett, and Trpkov (2014) demonstrated that the Internet is gaining popularity among health consumers and is becoming the first source of health information.

1.4.1 Access to online health information and its use globally

It is important to know whether access to the Internet is available globally for health consumers. Internet World Stats (2017c) reported the number of worldwide Internet users on March 25, 2017. Those data are shown in Table 1.

Table 1: *Internet World Statistics by World Internet Usage and Population Statistics*

<u>World Regions</u>	<u>Population % (re: World)</u>	<u>Penetration Rate</u>
Africa	16.6%	27.7%
Asia	55.2%	45.2%
Europe	10.9%	77.4%
Latin America/Caribbean	8.6%	59.6%
Middle East	3.3%	56.7%
North America	4.8%	88.1%
Oceania/ Australia	0.5%	68.1%
WORLD TOTAL	100%	49.6%

Note: Reprinted from (Internet World Stats, 2017c), second table.

Penetration rate: proportion of people in the region who use the Internet.

The majority of Internet users live in Asia (55.2%); however, Asia's Internet penetration rate is only 45.2%. The Internet penetration rate mentioned in this table according to Internet World Stats (2017a) is "the percentage of the total population of a given country or region that uses the Internet" (section 1.3 Internet Penetration Rate). The highest Internet penetration rate is 96.9% in the Falklands Islands, and Iceland has the second highest rate with 96.5% (Internet World Stats, 2017b).

The determinants of different penetration rates globally was studied by Chinn and Fairlie (2007). The researchers examined 161 countries across several variables: economic (trade openness, illiteracy, years of schooling), demographic (youth and aged dependency ratios and urbanisation rate), infrastructure (density of telephones, consumption of electricity),

regulatory quality and telecommunication pricing measures. The results illustrated that all the above-mentioned variables significantly affect the computer and Internet use among people, except for trade openness and telecommunication pricing measures. The researchers also suggested that “public investment in human capital, telecommunications infrastructure, and the regulatory infrastructure may mitigate the gap in PC and Internet use” (p.16). In addition to differing Internet penetration rates, there may also be different rates of Internet use for health information.

In the United States of America, on an average day, approximately 8 million individuals use the Internet to search for health-related problems (Fox, 2012). A cross-sectional study conducted by Moreland, French, and Cumming (2015) to find the prevalence of Internet usage in 10 primary care centres in a rural community of Scotland demonstrated that 68.4% of patients coming to the facilities had used the Internet to search for health information. The researchers made a clear conclusion in this study about the increasing number of customers using computers for their health-related queries, which should be studied further in terms of the relationship between the medical professional and the patient (Moreland et al., 2015).

1.4.2 Access to online health information and its use in India

The latest report provided by the International Telecommunications Union (2016) published on 15 September 2016 mentions that “India has overtaken the United States to become the second largest Internet market, with 333 million users, trailing China’s 721 million” (para. 1). According to this report, India has also become the second-largest smartphone market with approximately 260 million broadband mobile subscriptions. Another report published by the International Telecommunications Union (2016) illustrates that mobile broadband use has increased by a factor of nearly 2.5 in India over the last two years. There is an opportunity to expand broadband connectivity in India making it the “next big frontier” of the world (International Telecommunications Union, 2016). In order to achieve this, the

government of India is executing a programme called “Digital India”, aiming to connect rural councils via broadband (International Telecommunications Union, 2016). In India, the Internet penetration rate is 28.3% (Internet World Stats, 2015) which is very low compared to other countries. It is important to note that these data include mobile Internet usage, therefore this Internet penetration rate is not definitive.

An extensive search of literature to find the use of the Internet for health information in India resulted in only one study, that conducted by Akerkar, Kanitkar, and Bichile (2005). This study was executed in a private, tertiary care hospital located in Mumbai city and a questionnaire was given to 880 consecutive adults aged 18–70 years visiting the hospital. The results showed that one out of four patients visiting the hospital used the Internet to search for health information. This study also found that the patients using the Internet to obtain health-related information are satisfied with the quality of the information displayed on the webpages.

1.4.3 Access to online health information in Hindi speaking population

A news article published by Google (Google.org, 2016) shows that Google introduced an idea to translate medical information in Hindi and other Indian regional languages. According to a news article published by Press Trust of India (2017), it was anticipated that 536 million Internet users will use regional languages for searching and using the Internet by 2021. Moreover, Hindi Internet users (201 million) will outnumber English Internet users by 2021 in India. These regional languages are not only used for online chats and entertainment, but are used for different payments methods such as online banking (Press Trust of India, 2017). The factors affecting the growth of Internet usage in Indian languages depend on the availability of good quality Indian language webpages, high speed Internet and digital literacy (Press Trust of India, 2017). Although there are webpages available for hearing-related information in Hindi, there is no academic literature present at the moment identifying the usage of Hindi webpages for searching for health information.

1.5 Health literacy

1.5.1 Health literacy globally

The ability of Internet users to understand online health information depends, in part, on their health literacy. Health literacy has gained the attention of researchers who aim to reduce the gap between consumers and health care providers in terms of knowledge about health information. Sørensen et al. (2012) conducted a study examining different aspects of health literacy to create a complete meaning and a conceptual model describing the broad research-based dimensions of health literacy. The resulting definition from this study was “Health literacy is linked to literacy and entails people's knowledge, motivation and competences to access, understand, appraise, and apply health information in order to make judgements and take decisions in everyday life concerning healthcare, disease prevention and health promotion to maintain or improve quality of life during the life course” (p. 3).

To determine the impact of health literacy on the access of health information and its use leading to a poor health outcome, a systematic review was performed in the United States of America (Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011). The researchers concluded that poor health outcomes and usage of medical care facilities is significantly associated with low health literacy rates which puts them at greater risk for poorer access to care and poorer health outcomes. The poor health outcomes mean a rise in hospitalisation rates, inability to accurately read the labels on medication slips and inability to consume prescription drugs correctly. Another systematic review performed by DeWalt, Berkman, Sheridan, Lohr, and Pignone (2004) found that people reading at a low level are 1.5 to 3 times more prone to poor health outcomes than people reading at a higher level. Therefore, health literacy depends on the reading ability of the person. One major factor affecting reading ability is the reading grade level (RGL), obtained by students in the school education system, of the written material. Doak (1996) considered a person to be literate if they can read at an RGL of 5 or higher. Data

from a survey by the National Centre of Education in 2003 found that 22% of the American adult population (aged 16 years and older) had a basic level of health literacy and could effectively read material at the 8th or 9th RGL. In addition, it was also observed that 1 in 5 adults could only read at the RGL of 5 or below (National Center for Education Statistics, 2003).

A study performed by van der Heide et al. (2013) in the Dutch population provided strong evidence illustrating the association between health literacy and education. The researchers suggested that various means used to decrease disparities in health outcomes can benefit if more attention is given to health literacy. It means that if health information can be easily read and understood by users, there are higher chances of better health outcomes. From the research evidence, it is clear that one way to improve health outcomes is to ensure that the health information available (online or in any other medium) can be read and understood easily by everyone. To achieve this aim, researchers have developed literacy guidelines to construct written health information for people with different academic or health backgrounds (Sørensen et al., 2012).

1.5.2 Health literacy in India

The level of literacy in India was surveyed by Ministry of Home Affairs (2001) and literacy was mentioned as “an important input in overall development of individuals enabling them to comprehend their social, political and cultural environment better and respond to it appropriately” (para. 1) . In this survey, 2 million people were allotted to gather the information by visiting each household in India and each member of the household was asked about their level of education. The overall literacy rate was found to be 64.8% (male: 75.3% and female: 53.7%). Literacy rate in this survey (Ministry of Home Affairs, 2001) was defined as “the percentage of the population of an area, aged seven years or above, who could read and write with understanding” (para. 2) .

No language was specified in this survey. In cities, a higher literacy rate was documented in both males and females and the gender gap reduced to 13 percentage points. Adult literacy and youth literacy rates (15–24 years) were also determined in another survey conducted by Ministry of Home Affairs (2011). Adult (24+) literacy rate increased to 74.04% and youth literacy (15–24) rate to 90%. To date, there are no data available determining the level of health literacy in India. However, the overall low literacy level suggests a greater need to take steps to construct health material suitable for the Indian population.

1.6 Quality of online health information

Although the Internet has become a convenient tool to assess online health information for consumers, challenges regarding the quality of material available online still persist (Fiksdal et al., 2014). Due to an absence of regulatory guidelines for quality of online material, it is possible that online information can be of poor quality, potentially misleading and highly variable.

Studies have identified the variability of online medical information on various topics. Eysenbach, Powell, Kuss, and Sa (2002), in a systematic review of published articles identifying the quality of online health care information, concluded that the health care information available on the Internet is not adequate; this problem needs to be addressed.

1.6.1 Readability

DuBay (2004) defines readability as “what makes some texts easier to read than others” (p. 3). The most comprehensive definition given by Dale and Chall (1949) is: “The sum total (including all the interactions) of all those elements within a given piece of printed material that affect the success a group of readers have with it. The success is the extent to which they understand it, read it at an optimal speed, and find it interesting” (p. 23). Hence, the understanding of a text depends on the level of its readability. If the readability level is higher, there is more chance that readers will not be able to read and understand easily, and readers

may misinterpret the text (DuBay, 2004). Therefore, if the text of online health information has a satisfactory level of readability, the consumers will be able to understand the information.

RGL, according to Black (1910) is the “number assigned to the level of how complex the reading material is. It equals a level of schooling. For instance, a level of seven means the reading material is suited for a student in seventh grade” (para. 1). Readability and RGL can be affected by the style of text such as length of a sentence, frequency of polysyllabic words and jargon (Laplante-Lévesque et al., 2012). Researchers have attempted to identify an appropriate RGL for health material to be understood by the majority of the population. For instance, Weiss and Natl Work Grp Literacy (1998) provided a specific RGL of 5 or 6 in order for health material to be understood by the majority of the U.S. population. Doak (1996) also recommended an RGL of 5 or 6 for any type of health information and if the RGL is between 6 and 8, the material can be considered as adequate, whereas above 9, the material can be considered unsuitable.

A systematic literature review was performed by Laplante-Lévesque and Thorén (2015) to determine the RGL of online hearing health information. The authors identified 78 articles in their review, and the mean RGL of online hearing-related information ranged from 9 to over 14. The conclusion of this systematic review proposed by Laplante-Lévesque and Thorén (2015) was that the hearing information available on the Internet has poor readability and there is a great need to take some initiatives to address this issue.

In a study conducted by Laplante-Lévesque et al. (2012), the readability of hearing-related information was assessed; it was obtained by using the Google search engine after entering the terms “hearing aids” and “hearing impairment”. The authors used the English language Google domains of Canada, India, Australia, United Kingdom and United States of America. The first 10 webpages were retrieved from each search and three readability formulas were used to calculate the RGL. The researchers observed that the majority of the text of

webpages available from the Google search was at an RGL of 9 or more, indicating that the information is hard to understand for many consumers.

Recently Hsu (2017) examined the readability of online hearing-related webpages in Traditional Chinese. In his study, he recruited 39 native Chinese speakers who read Traditional Chinese and asked them to give some search terms most commonly used to search for any type of hearing-related information. The most commonly used search terms were entered in a search engine and the first ten webpages obtained after entering each search term were evaluated. After applying inclusion and exclusion criteria, 31 webpages were analysed by using two readability formulas. Hsu found that 25% of the webpages had an RGL of more than 6 by using the CRIE1.0 formula, and 81% of the webpages had an RGL of more than 6 when the Jing formula was used. Moreover, after sorting webpages according to type of organisations, no significant difference in RGL was found.

1.6.2 Other measures for quality control

Bernstam, Shelton, Walji, and Meric-Bernstam (2005) conducted a study to identify tools to analyse the quality of the health care information provided on the Internet for health consumers. The authors found 273 different tools to assess the quality and analysed them depending on various criteria such as “number of elements, availability, readability and objectivity” (p. 13). Out of these 273 tools, only 80 tools were available to the general public and 24 tools had inadequate elements (particulars that a consumer has to identify to assess the quality of a website). Only seven tools contained adequate particulars. With this extensive research, it was clear that only a few quality-rating tools can be practically utilised by consumers.

The Health On the Net Code of Conduct (HON code) is a globally accepted and a well-known third-party certification tool (Boyer, Dolamic, Ranasinghe, & Baujard, 2015) developed by a nongovernmental Swiss organisation enabling Web developers. It was created for the

benefit of customers, medical care professionals and website publishers. According to the World Health Organization (2011) , “the presence of the HON code logo signifies to patients and providers that the site adheres to certain principles and has undergone HON’s certification process. The eight principles governing the HON code are: authority; complementarity; confidentiality; attribution; justifiability; transparency; financial disclosure; and advertising” (p. 6, para 2). The sites that satisfy these principles are given a seal by the HON foundation which links these webpages to the HON website displaying their performance. This certification process is voluntary and is performed by a review committee established by the HON foundation. After the first certification, the site is monitored periodically. If website developers want to make any changes, they need to notify the HON foundation. HON certification was free of charge in the early stages (1996), but now website owners need to pay some fee for the approval. It is available in over 30 languages (Boyer et al., 2015). Webpages can display a certification badge after getting approval by the HON foundation (Health On the Net Foundation, 2010).

Another important aspect relating to the assessment of the quality of health-related webpages is the origin or type of the organisation creating the website. Generally, organisations are of three types: commercial/for profit, nonprofit, and governmental. When there is a commercial motive behind the webpages, the website owners can publish a website in favour of their product, which can be dangerous (Cline & Haynes, 2001). There is a potential for readability of texts to be easier for the commercial webpages than for nonprofit and governmental webpages because they want to make a profit (Cline & Haynes, 2001). This can potentially directly impact the quality of online health material, and therefore needs to be addressed.

1.6.3 Cloze test

The cloze test was published (Taylor, 1953) after identifying several difficulties with the standard readability formulas such as those of Dale-Chall and Flesch (DuBay, 2004). Taylor argued that words per se are not the measure of reading difficulty, but rather it is their relationship with each other. Therefore, Taylor proposed a new procedure called the cloze test to measure a person's understanding of a text. The inherent theory of this test is the ability of a person to fill in the missing words of a text, based on the surrounding context, thereby providing closure to the meaning of the text.

a. Procedure of the cloze test

In the cloze test, every n th word (commonly every fifth word) of a sample of text is deleted and replaced with a blank. Respondents fill in each blank with the word they think best completes the text (DuBay, 2004). After deleting every n th word, the replaced blank should be of a similar length to the original word (DuBay, 2004). The results obtained, the cloze score, is reported as the percentage of blanks correctly filled in. Lower scores indicate more difficulty in understanding the text. The rationale behind this test is that the respondent must be susceptible to semantic and syntactic restrictions in each context to fill in the blanks (Shahnazari-Dorcheh, Roshan, & Hesabi, 2012). The length of the cloze test studied by Shahnazari-Dorcheh et al. (2012) found that the length of the cloze test depends on the proficiency level. The authors concluded that at the advanced proficiency level, a long test including 50 empty spaces is reliable and a short test including 20 to 25 spaces is more reliable at beginner and intermediate proficiency levels.

b. Scoring of cloze test

Scoring of the cloze test can be performed by different techniques such as the “exact-word” scoring technique, quick performance measurement and feedback technique, and the three-stage scoring hierarchy for partial credit. In the exact-word scoring technique, the score is awarded only if the exact word is entered into the blank space (Clausing & Senko, 1978). In the quick performance measurement and feedback technique, various types of marking are possible. For the exact match of the word, two points are given and one point is given for a partially acceptable response, and zero points are awarded for a completely different response (judged by the grader). The three-stage scoring hierarchy for partial credit technique scores depend on three areas: context, orthography (standard rules for representing the language in the written form) and grammar (Clausing & Senko, 1978).

According to DuBay (2004), it is hard to complete more than 65% of the deleted words even by readers with advanced reading skills. Assisted reading texts need a cloze score of 35% or more. A higher cloze score is needed for unassisted reading text. Cloze scores can be compared with the answers of a multiple-choice questionnaire, as shown in Table 2.

Table 2: *Cloze and multiple-choice scores comparison*

<u>Purpose</u>	<u>Cloze</u>	<u>Multiple-Choice</u>
Unassisted reading	50–60%	70–80%
Instructional, assisted reading	35–50%	50–60%
Frustration level	Below 35%	Below 50%
<i>Note:</i> Reprinted from DuBay (2004), p. 27.		

1.7 Readability analyses

Readability analysis is carried out to find the RGL of any type of text information. It can be calculated by computational and noncomputational methods. Computational methods use readability formulas to assess readability and hence are easy to use and reliable. In the literature, a number of readability formulas are available to calculate the readability of texts in different languages. The first readability formula, Flesch Reading Ease (FRE) Score in the English language, for measuring adult reading material, was published by Flesch in 1942 (DuBay, 2004). The Flesch–Kincaid (F–K; Kincaid, Fishburne Jr, Rogers, and Chissom (1975) and Simple Measure of Gobbledygook, SMOG (McLaughlin, 1969) are commonly used readability formulas in English. Emphasis on easy understanding of text and quantification are the two main characteristics of these formulas, which means that the readability formulas can assess the text difficulty objectively and quantifiably (Bailin & Grafstein, 2001). In addition, the RGL is obtained as a numerical score that can give the user a sense of the level of text difficulty (Bailin & Grafstein, 2001). Due to their easy implementation, many schools and libraries use these formulas to evaluate the readability of textbooks (Carter, 2000).

Noncomputational methods were developed to measure comprehension of the text and, unlike computational methods, they require actual human perception of the text (Taylor, 1953). An example of a noncomputational method is the cloze test which has been explained in the previous section. They are more time consuming and less practical for large-scale analysis due to the requirement of finding participants. In my study, I will use both measures to analyse readability of text.

1.7.1 Readability analysis in Hindi

Readability formulas have been designed for different languages based on the studies that developed English readability formulas. Sinha et al. (2012) conducted a study and

demonstrated “the inadequacy and the consequent inapplicability of some of the popular readability metrics in English to Hindi” (p. 1141). In a similar study, the authors proposed two readability formulas combining all the salient features of Hindi (Sinha et al., 2012). To date, there are no other readability formulas in Hindi available in the literature.

1.7.2 Features of Hindi language

Kachru (2006) mentioned that Hindi is spoken by 300 million people in the northern part of India as their first or second language. Hindi, which originated from Sanskrit, became an official language of India in 1965 (UCLA Language, 2014). Medieval Apabhramsa and Middle Indo-Aryan languages provided a pathway to the origin of Hindi from Sanskrit. Old Hindi appeared around AD 500 and Modern Hindi was developed in the period from the thirteenth to the eighteenth centuries (UCLA Language, 2014). Hindi is written in the *Devanagari script*, which is written from left to right. Sanskrit is the origin of Devanagari script and other key languages like Nepali, Sindhi and Marathi (UCLA Language, 2014). Hindi is the second most common language spoken in the world in terms of the numbers, after Chinese (UCLA Language, 2014).

The Hindi language consists of 11 vowels called *swar* and 33 simple consonants called *vyanjans* (Kellogg, 1876). Consonants are joined together to make a special conjunct called a *jukta-akshar* and vowels have a special characteristic in Hindi that they can be written independently or by diacritical marks called *matras* (Kellogg, 1876). In Hindi, *halant* is also used to remove *matras* and is also called a vowel omission sign (Kellogg, 1876).

1.7.3 Differences between Hindi and English

There are many differences between Hindi and English which are not necessary to mention in the context of this study. However, there are some differences which will help in better understanding this study:

1. Grammatical inflection defined by Vikram (2013) is “combination of a word stem with a grammatical morpheme, usually resulting in a word of the same class as the original stem” (p. 2). The inflection in English for nouns can reach up to 7 or 8 while it can reach up to 40 or more in Hindi (Vikram, 2013).
2. In the English language, verbs never change according to gender but in Hindi, verbs do change according to gender (Kellogg, 1876).
3. Polysyllabic words are not necessarily considered “hard” in Hindi, but they are considered “hard” in the English language. For instance: In Hindi, a word like “ बनवाऊँगी” (“banvaoongi”) has 4 syllables but is very easy to read and understand (Sharma, 2012).
4. Hindi is a head-final language with *postpositional case marking*. Davison (1999) mentioned that “some postpositions are associated with grammatical function, some with specific roles associated with the meaning of the verb” (para. 2).
5. In English, letters are written one after the other but in Hindi, vowel marks (*matrayen*) are used to make syllables when applied over consonants (Kellogg, 1876).
6. Unlike English, Hindi is a phonetic language; which means it is spoken the same way as it is written (Sharma, 2012).
7. In Hindi, jukta-akshars are commonly used, which means there is more than one representation of a given conjunct (refer to Section 1.7.1).

1.7.4 Readability computational measures available in Hindi

Readability measures available for English cannot be used directly to find the level of readability in the Hindi language (Sinha et al., 2012). Sinha et al. (2012) mentioned the reason for the inapplicability of English readability is that Hindi and English have a significant difference in basic structure (refer Section 1.7.3). As compared with the extensive readability research carried out in other languages, fewer efforts have been made to develop computational readability measures for Hindi texts. To develop a computational readability measure for Hindi,

Sinha, Dasgupta, and Basu (2014,) designed an online-computational tool which could be used by researchers to assess the readability of various texts.

A Graphical User Interface (GUI) was developed by using Java script (ORACLE, 2015). The designed application (Automated Readability Processing Tool) could be run on Windows, MacOS X , Linux and other platforms (Sharma, 2012). To run this tool, the text to be examined for readability should be clean, which means all the symbols and spaces are removed. This file will work as an input file for the computational tool. After entering the input file into this tool, text will be parsed and the tool will display all the textual factors required to calculate readability such as word length, sentence length, syllable count, frequency of words, count of jukta-akshars and polysyllabic words (Sharma, 2012). An example of the interface is shown in Figure 1.

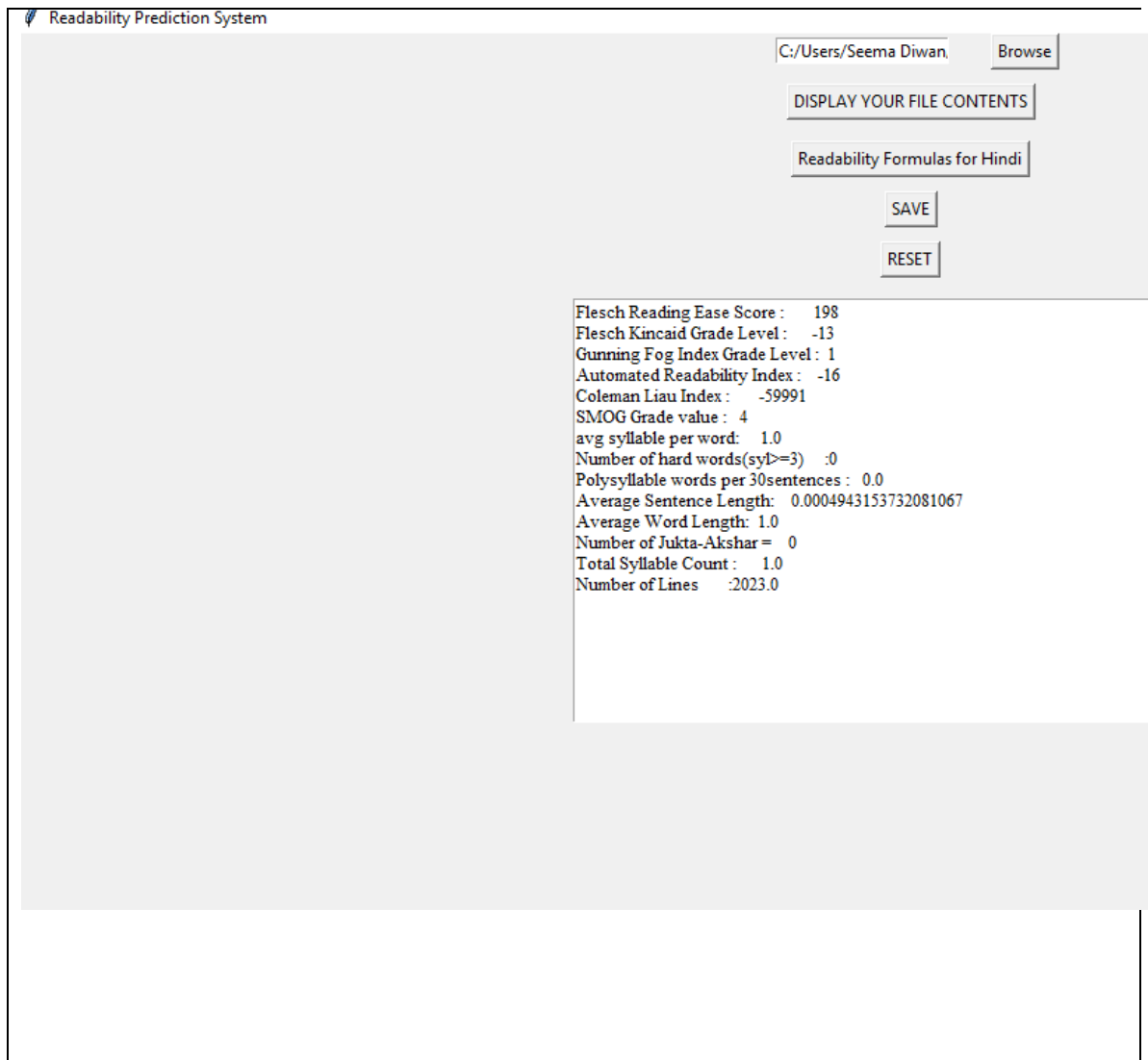


Figure 1: *Graphical User Interface (GUI) of the computational tool*

➤ Hindi structural parameters of a text

The following structural parameters were used to compute readability in Hindi in this study (Sinha et al., 2012)

c. Average sentence length (ASL)

Average sentence length is the total of the sentence lengths divided by the total number of sentences. Sentence length is obtained by adding the number of words in a sentence after

omission of all spaces or any kind of symbol. Each sentence in a text is separated by a *purnaviram* (equivalent to full stop or period in English) or a punctuation mark.

d. Average word length (AWL)

AWL is the total length of all the words divided by the number of words. The word length is calculated by splitting a word into akshars, and matras are omitted, but the inherent matra of “a” (अ) and akshars with halant are included.

Example: काल्पनिक (“Kalpanik”) = क + ा + ल् + प + न + ि + क AWL =5

e. Syllable count, polysyllabic words (PSW)

The number of syllables are counted by adding the number of free vowels, number of matras and half schwas. Free vowels are calculated by identifying the independent akshars and schwas is a vyanjan without matra or halant.

Example: मन्त्र (“Mantar”) = मं + त्र + र, Number of free vowels= 1 (मं), Number of matras = 0,

Number of schwas =2 (त्र, र); Syllable count= $1+0+2/2 = 2$. PSW are the words containing more than 2 syllables.

f. PSW30

PSW per 30 sentences (normalised for 30 sentences) is calculated by dividing the total number of PSW in a text by the total number of sentences and multiplying it by 30.

g. Average number of syllables per word (ASW)

ASW is calculated by dividing total syllable count by total number of words.

h. Jukta-akshars (JUK) or consonant conjuncts

A JUK is an akshar (consonant) clustered with other another akshar (consonant). A JUK is counted as such when an akshar with halant comes after another akshar. If an akshar with halant is not followed by another akshar, it is not identified as a JUK.

Example: संकल्प (“Sankalp”) = स + ं + क + ल + ् + प JUK = 2

➤ **Readability formulas in Hindi**

Readability formulas in Hindi were developed by Sinha et al. (2012) in a series of experiments to identify the salient features of Hindi affecting readability. The authors proposed two different readability models for Hindi and tested them with a new data set. The results obtained in this manner validated the models. The study helped in the development of two new readability measures or formulas: RH1 and RH2 for Hindi readability. AWL, PSW, JUK and PSW30 (number of polysyllabic words per 30 sentences) were acknowledged as the major factors affecting Hindi text readability. In previous studies (Agnihotri & Khanna, 1991; Bhagoliwal, 1961), English readability formulas were applied to compute the readability of Hindi textbooks. However, they did not provide a complete model for calculating Hindi readability like those of Sinha et al. (2012). Therefore, I decided to use RH1 and RH2 in my study, which were:

$$RH1 = -2.34 + 2.14 * AWL + 0.01 * PSW$$

$$RH2 = 0.211 + 1.37 * AWL + 0.005 * JUK$$

1.8 Study rationale

The above research into literature about the readability of online text about hearing health information clearly illustrates that it has a high RGL (more than 6), making it hard to read and therefore reducing its usefulness. With the intention of making online hearing-related information be useful for health consumers, research should be done to identify the RGL of online material in different languages. Previous studies have only focused on identifying the readability level of online hearing information in English (Atcherson et al., 2014; Laplante-Lévesque & Thorén, 2015). To date, there is no research available that has investigated the readability of hearing-related information available on the Internet in the Hindi language. Given the large population of Hindi speakers in the world, there is a clear demand for the assessment of text readability available on Hindi webpages.

The present study intends to replicate the research performed by Hsu (2017), who assessed the readability of hearing-related webpages in traditional Chinese. In addition, my study will also use cloze tests to assess the readability of these webpages, and HON certification to assess the content quality.

1.9 Aims and hypothesis

In regard to assessing readability by using readability formulas (Part 1), the research question of this part of the study is: What is the mean readability (RGL) of online hearing-related webpages in Hindi?

In regard to assessing the readability of online hearing-related information by using a cloze test (Part 2), the research question of this part of the study is: Is there a significant difference in cloze scores for the webpages with the highest and lowest mean RGL?

The planned alternative (research) hypotheses were:

Hypothesis 1: The mean readability score (RGL) of hearing-related webpages published in Hindi is significantly different from 6.

Hypothesis 2: The mean readability score (RGL) of commercial webpages will be lower than the mean RGL for government or nonprofit organisation webpages.

Hypothesis 3: There is a significantly even distribution of HON certification across the types of organisations.

Hypothesis 4: There is a significant difference between the readability score (RGL) calculated by RH1 and RH2 formulas.

Hypothesis 5: There is a significant correlation between the readability scores (RGL) calculated by RH1 and RH2.

Hypothesis 6: There is a significant difference between the Cloze scores and RGL of the webpages.

2. CHAPTER TWO: METHODS

This study was conducted in parts at the University of Canterbury. In Part 1, the readability of online hearing health-related information available in Hindi was calculated by using two readability formulas, RH1 and RH2 (Sinha et al., 2012). In Part 2, a cloze test (Taylor, 1953) was conducted to assess the quality of the health information of the webpages with highest and lowest mean RGL.

2.1 Part One (Readability)

2.1.1 Identification of search terms

The identification of key words for the Internet search was done by recruiting a group of people who spoke Hindi as their first language. The 25 informants were identified through Facebook and through personal links. They were asked the following question in Hindi: “If you realise that you are having hearing problems or difficulties and you want to look for general information about a hearing problem and its treatment, which search terms will you try on the Google search engine? Please do not hesitate to mention as many as you like.” Two informants declined to identify search terms because they said they did not use the Hindi language for Internet searches. The search phrases or keywords identified by the informants were: कान की समस्या (ear problems), कान की मशीन (hearing aid), कान की परेशानी (ear troubles), कान में दर्द (ear pain), सुनने में परेशानी (hearing trouble), कम सुनाई देना (hearing impairment), and बहरेपन के लक्षण (signs of deafness). Out of these seven phrases, five search phrases were selected on the basis of their use by two or more informants; these were: सुनने में परेशानी (hearing trouble), कान की समस्या (ear problems), कान की मशीन (hearing aid), बहरेपन के लक्षण (signs of deafness), कम सुनाई देना (hearing impairment).

2.1.2 Internet search

A 13-inch MacBook Pro having OS X El Capitan operating system version 10.11.6 was used for performing the online search. Google Chrome browser (version 58.0.3029.110) was used for the search and for checking HON certification by using this link: <http://www.hon.ch/HONsearch/Patients/index.html> (Baujard, Boyer, & Geissbühler, 2010).

The country-coded Top Level Domain (ccTLD) can be defined as the most commonly used domain for Internet searches in any country. The ccTLD is Google.co.in for India and it is available in Hindi. Each search phrase was entered one by one into Google India (google.co.in) Hindi version (named Google Bharat). The uniform resource locators (URLs) of the first ten webpage results obtained after entering each search phrase were recorded. The webpages were used as units for analysis, so one webpage was treated as one unit or participant.

2.1.3 Inclusion and exclusion criteria

The inclusion criteria for the selection of webpages were: (1) must be in Hindi, (2) must contain hearing or hearing impairment-related information, (3) must be available to the public and (4) must contain information about the organisation hosting the webpage. There was no information about the organisation hosting the webpage on some webpages. In those cases, the information was collected by searching through a separate Google search.

The exclusion criteria for the selection of webpages were: (1) a Google-identified advertisement, (2) a video, (3) a directory listing, (4) contains less than 100 words, (5) contains information on tinnitus, otitis media, tumours, vestibular disorders, (7) contains images only.

2.1.4 Webpage information

The search was conducted on 12/04/2017 at the University of Canterbury and was started by entering five identified search terms in Google Bharat (ccTLD of India). The first ten webpages resulting from each search were recorded during the process. The previous literature has established that consumers tend to access the first page only while searching

their key terms (Eysenbach & Köhler, 2002). Fifty webpages were provided after entering five search phrases into one search engine and these webpages were assessed against the inclusion and exclusion criteria.

The origin of each webpage was recorded in the form of type of organisation (commercial/for profit, nonprofit, and governmental). The type and country location of each organisation were retrieved from the “about us” (or similar) section on each webpage. The URL of each webpage was copied and pasted into a Microsoft Excel file. HON code certification of webpages was assessed by using this link: <http://www.hon.ch/HONsearch/Patients/index.html> (Baujard et al., 2010) using Google Chrome. If the last date of update on the webpage was mentioned, it was recorded. No special tool was used to access the last date of update.

2.1.5 Selection of a paragraph

The relevant content in Hindi on the webpage was copied and pasted into a Microsoft Word document. Each sentence was numbered and a random generator in an Excel spreadsheet was used to select the first sentence for the analysis. The portion of the text to be explored was composed of the first sentence starting from that random number and the next 100 words/characters, confirming that the last sentence was a full sentence. In this manner, the selected paragraph consisted of at least 100 words or characters.

2.1.6 Readability computational tool

A readability computational tool was downloaded on a computer with Windows 10 version 1607. The selected paragraph was entered in this tool and AWL, PSW) and JUK (juka akshar) were retrieved (Sinha et al., 2012). The results obtained by this analysis of the computational tool were recorded in an Excel spreadsheet. The RH1 and RH2 formulas (Sinha et al., 2012) were used to calculate the RGL of each webpage with the help of a calculator.

2.2 Part Two (Cloze test)

In Part 2, the cloze test was performed to obtain more information about the readability of online hearing-related webpages in Hindi.

2.2.1 Ethics Approval

The approval was gained from the Human Ethics Committee of University of Canterbury for running this procedure on 23 May 2017 (Appendix A.1). The cloze test was carried out according to the approved ethics application.

2.2.2 Selection of paragraphs

The paragraphs on the webpages that resulted in the lowest (easiest to read) and the highest (most difficult to read) readability level were identified and were used in this procedure. In the identified paragraphs, every fifth word was replaced by a blank.

2.2.3 Participants

Ten participants were recruited over a two-week period according to the inclusion and exclusion criteria detailed in Table 3. The exclusion criteria attempted to ensure that the knowledge about hearing health would not influence the responses given by participants.

Table 3: *Inclusion criteria and exclusion criteria for participants*

Inclusion Criteria	Exclusion Criteria
1. Participants must be aged 18 years and over	Participants must not have any expertise in the hearing health industry
2. Participants must be native Hindi speakers of any gender	

2.2.4 Procedures

Participants were identified from a list of Facebook friends. A message was posted on my Facebook timeline which mentioned the topic of research only. The post displayed my email address and my supervisor's email address, and the interested person could email me or my supervisor, and the information sheet was sent to them. In this way, ten interested participants were sent the following material: (1) a demographic questionnaire (Appendix A.2), (2) information sheet (Appendix A.3), (3) consent form (Appendix A.4), (4) Cloze test (Appendix A.5) and a postage-paid return envelope. Participants were instructed to sign the consent form after reading the information sheet fully, and to fill in the demographic questionnaire. Finally, they were asked to complete the cloze test. All the participants returned the appendices and none of them withdrew from the study.

2.2.5 Measures

- **Demographic questionnaire (Appendix A.2)**

The demographic questionnaire was designed to get basic information about the participants in the study. The items on this questionnaire were: ID number (given by me), age, gender, years of education completed, relationship status and their ethnic group.

- **Cloze procedure (Appendix A.5)**

Two paragraphs in Hindi (selected after readability analysis) having blanks were displayed on this sheet. The participant needed to fill in the blanks, using only the information provided in the paragraph. If the task was found difficult, participants were encouraged to make a guess. In this study, the interpretation of the cloze scores will be as follows (Taylor, 1953):

0–39% = not understood

40–59% = usable, may require additional information

60–100% = understood

2.2.6 Planned statistical analysis

The data obtained by the above-mentioned procedure was statistically analysed by IBM SPSS statistics version 24. Descriptive analyses were performed to answer the first, second and third hypotheses. Nonparametric Spearman's rho was performed to answer the fourth hypothesis. A summary of the analyses for each hypothesis is in Table 4.

Table 4: *Planned statistical analyses of each hypothesis*

Hypothesis	Statistical Analysis
1. The mean readability score (RGL) of hearing-related webpages published in Hindi is significantly different from 6	Descriptive analyses, One-sample t test
2. The mean readability score (RGL) of commercial webpages will be significantly different from the mean RGL for government or nonprofit organisation webpages.	Descriptive analyses, Analysis of Variance (ANOVA)
3. There is a significantly even distribution of HON certification across the types of organisations.	Chi square test and descriptive analyses
4. There is a significant difference between the readability score (RGL) calculated by RH1 and RH2 formulas.	Related sample t test
5. There is a significant correlation between the readability scores (RGL) calculated by RH1 and RH2.	Pearson correlation
6. There is a significant difference between the cloze scores and RGL of the webpages.	Related sample t test

3. CHAPTER THREE: RESULTS

3.1 Overview

Readability analyses and cloze tests were performed for 25 webpages (out of 50 webpages: 5 search terms + 10 results) which were obtained after removing duplicates and applying inclusion and exclusion criteria. In this section, the results are explained in two parts: readability and cloze test.

3.2 Part 1 (Readability)

The RGLs of 21 webpages were obtained with the help of a computational tool. The computational tool could not assess the content of the remaining four webpages, therefore the RGLs of these webpages were calculated manually. The mean RGL was also calculated in the Excel spreadsheet after obtaining RGLs of the webpages by using the RH1 and RH2 formulas. In addition, the date of the last update was retrieved from 16 webpages and no webpage was HON certified. The content of two webpages could not be copied; therefore, the content was typed manually for readability analysis. Type of organisation and country of origin were easily retrieved and recorded. SPSS version 24 (IBM, 2016) software for MacBook Pro was used for the statistical analyses. Descriptive statistics obtained from RH1 and RH2, and the mean RGL of the webpages are illustrated in Table 5 and Figure 2. The minimum RGL of hearing-related webpages published in Hindi obtained by RH1 was 3.03 and by RH2 it was 3.12. The maximum RGL by RH1 was 11.40 and by RH2 was 5.74.

The statistics values of skewness and kurtosis indicated that the data was not normally distributed. This necessitated the use of nonparametric statistics.

Table 5: Means, standard deviations, minimum, maximum values, and sample sizes for each readability formula

	<i>N</i>	Min	Max	Mean	SD
RH1	25	3.03	11.40	5.95	2.03
RH2	25	3.12	5.74	4.72	.78
Mean	25	3.08	8.42	5.33	1.3

Note: *N* = Sample Size, Min = Minimum value, Max = Maximum value and
SD = standard deviation

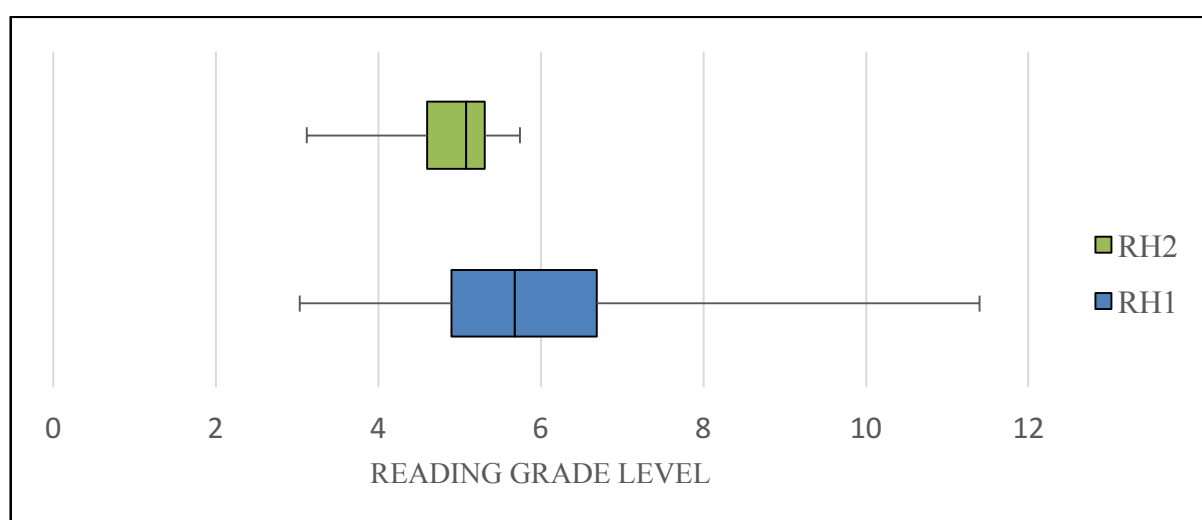


Figure 2: Box Plot of Reading Grade Level of Hindi webpages. RH1 = Readability Hindi 1.
RH2 = Readability Hindi 2.

3.2.1 Hypothesis 1

It was hypothesised that the mean RGL of hearing-related webpages published in Hindi is significantly different from 6. A one-sample t test was not performed because the data did not meet the assumption of normality. Instead a one-sample Kolmogorov-Smirnov test was used to test the hypothesis. The results showed that the mean RGL of hearing-related webpages

published in Hindi is 5.33, which is not significantly different from the recommended value of 6 ($p = .200$).

3.2.2 Hypothesis 2

Because the data did not meet the assumption of normality, a Kruskal–Wallis test was used to investigate whether the mean RGLs for webpages of the three origins (commercial/for profit, nonprofit, and governmental) were significantly different from each other. The summary of the results is shown in Table 6.

Table 6: *Kruskal–Wallis Test Results*

Type		<i>N</i>	Mean Rank
Mean	Government	2	23.00
	Nonprofit	2	18.00
	Commercial	21	11.57
Mean	Exact Sig.		.01
Kruskal–Wallis H value			5.407
Degrees of freedom (df)			2

There was a statistically significant difference between the mean RGL for webpages of different origins [$H(2) = 5.407, p = .01$]. A Mann–Whitney U test was performed to find which origins were contributing to these significant differences. The results indicated there was no significant difference between the mean RGL for governmental and nonprofit organisations [$U = 2.00, Z = .00, p = 1.00$]. Similarly, another Mann–Whitney U test demonstrated the RGL of nonprofit webpages was not significantly different from the RGL of commercial webpages [$U = 11.00, Z = -1.091, p = .332$]. However, a Mann–Whitney U test showed that the mean RGL

for governmental webpages was significantly higher than the mean RGL for commercial webpages [$U = 1.00$, $Z = -2.182$, $p = .016$].

3.2.3 Hypothesis 3

The third hypothesis regarding HON certification could not be tested because no website found in this study was HON certified.

3.2.4 Hypothesis 4

A Wilcoxon signed-rank test was conducted instead of a related samples t test to test Hypothesis 4 because the data were not normally distributed. The results showed that the mean RGL calculated by RH1 was significantly higher than the mean RGL calculated by RH2 [$Z = -4.157$, $p < .001$].

3.2.5 Hypothesis 5

A nonparametric Spearman's rho (r_s) was used instead of Pearson's correlation because the data did not meet the assumption of normality. There was a significant, positive correlation between the RGL values calculated by RH1 and RH2 ($r_s = .954$, $p < .001$)- Figure 3.

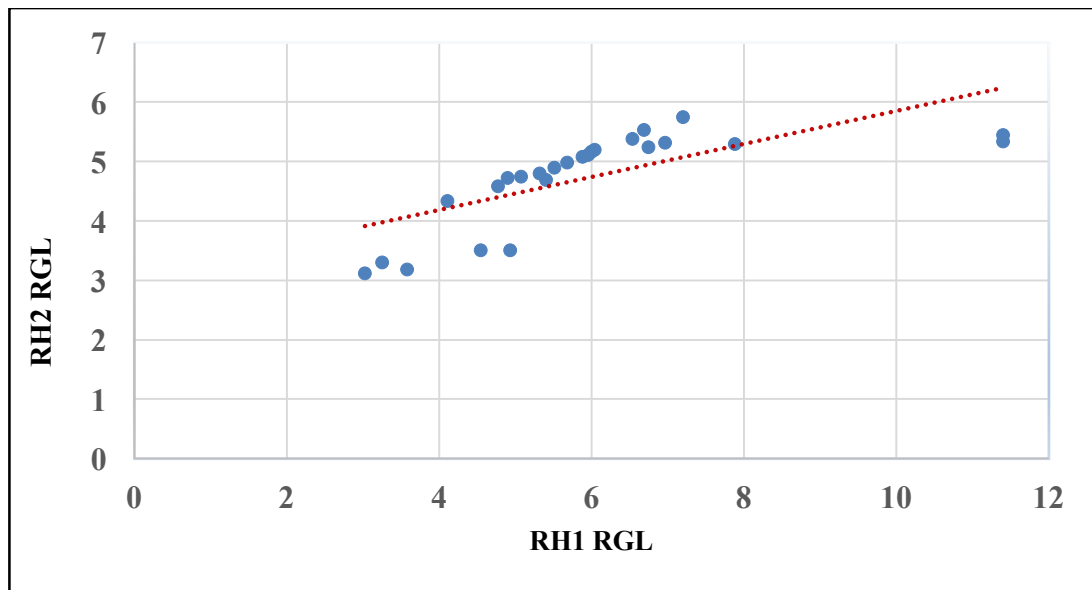


Figure 3: Relationship between RH1 and RH2 Formulas. Blue dots represent reading grade level (RGL) of each webpage and linear regression line is the red dotted line.

3.3 Part 2 (Cloze test)

3.3.1 Participants

Ten participants who showed interest in the research were from the northern part of India and native Hindi speakers. No participant dropped out of the study. All participants filled in the blanks of the text provided to them and returned them to me within a week. A further descriptive summary of participants is shown in Table 7.

Table 7: *Descriptive summary of participants*

Participant	Age	Gender	Years of Education	Relationship status	Ethnic Group
1	25	Male	15	Single	Indian
2	22	Female	20	Married	Indian
3	25	Female	22	Married	Indian
4	22	Male	19	Single	Indian
5	23	Male	20	Single	Indian
6	26	Male	20	Single	Indian
7	22	Male	18	Single	Indian
8	21	Male	15	Single	Indian
9	25	Male	18	Single	Indian
10	39	Male	15	Single	Indian

3.3.2 Hypothesis 6

Table 8 shows the cloze scores of the 10 participants from the two paragraphs. The table also shows the descriptive statistics for each paragraph.

Table 8: *Cloze score data obtained from participants and their descriptive summary*

Participants (min to max)	Cloze score on paragraph with maximum mean RGL (%)	Cloze score on paragraph with minimum mean RGL (%)
1	29.16 (Minimum)	34.7 (Minimum)
2	37.50	39.13
3	41.60	39.13
4	50.00	41.60
5	50.00	41.66
6	54.16	43.74
7	58.30	45.83
8	58.33	47.80
9	58.33	50.00
10	66.66 (Maximum)	50.00 (Maximum)
Mean	50.40	43.3
Standard deviation	11.36	5.06

A Wilcoxon signed-rank test was used because the data did not meet the assumption of normality. The Wilcoxon signed-rank test results indicated no significant difference between mean RGL and cloze scores [$Z = -1.779, p = .075$] with an effect size = 0.40.

4. CHAPTER FOUR: DISCUSSION

4.1 Introduction

In this study, 25 webpages were assessed for their readability using readability formulas and a cloze test. The Google Bharat user domain was used to retrieve these webpages using the five keywords सुनने में परेशानी (hearing trouble), कान की समस्या (ear problems), कान की मशीन (hearing aid), बहरेपन के लक्षण (signs of deafness), कम सुनाई देना (hearing impairment). The inclusion and exclusion criteria used in the study were similar to a study which was conducted recently to identify the reading grade level (RGL) of Chinese website information (Hsu, 2017). RH1 and RH2 readability formulas were used to calculate the RGL of online Hindi text available on hearing webpages. All webpages were arranged according to their origin, such as commercial, nonprofit or governmental. Commercial webpages inclined towards hearing aids and their features. On the other hand, nonprofit and governmental webpages were more focused towards the organisation and its purpose. For example, a government website provided information about various funding schemes in India and their candidacy. In the next sections, there is a discussion of the results of the RGL of online hearing-related information in Hindi by using RH1, RH2 readability formulas and the cloze test.

4.2 Relation to Literature and clinical significance

4.2.1 Readability of Online Hearing-related Health Information in Hindi

The RGL of online hearing-related information in Hindi did not exceed the recommended level for written health information. In this study, I applied the recommended RGL of 6 which exists in research performed for English health information (Doak, 1996; Weiss & Natl Work Grp Literacy, 1998). The mean RGL of the online hearing-related

webpages in Hindi was 5.33. In particular, the RH1 formula had a mean of 5.95 and RH2 a mean of 4.72. which suggests that Hindi online hearing-related information is not hard to read. These values indicate that the Hindi text available on these webpages is easy to read and there is no need to rewrite it in a simple language. Hence, the first hypothesis is not supported. This is the first study conducted to identify RGL of online hearing-related information in Hindi. However, similar types of studies conducted to find the RGL in English and Chinese did not show the same results. In a study completed by Hsu (2017), the mean RGL of Chinese webpages was 7.32 with a range of 4.16–12.25. In a systematic review (Laplante-Lévesque & Thorén, 2015), authors demonstrated that the mean RGL of online health-related information of webpages in English was 9 to over 14. The researchers concluded that there is enough evidence to say that the hearing-related information available on webpages in English has poor readability and this issue must be addressed immediately to provide maximum benefits to consumers from that information. However, my study did not achieve the same results. A possible explanation for this might be that the Hindi readability formulas lack special lexical attributes of text used in hearing-related information.

With respect to the second hypothesis, it was found that the RGLs of the webpages of government origin was significantly higher than those of commercial and nonprofit webpages, and no difference in RGL was observed between commercial and nonprofit ones. The high RGLs of the Chinese webpages of government origin obtained by Hsu (2017) support my findings. However, Hsu (2017) demonstrated a significantly higher RGL for nonprofit organisations than for commercial ones. My study did not support these findings and the small sample size of nonprofit organisation webpages should be taken into consideration before applying my findings clinically. The higher RGL of governmental organisation webpages compared with commercial ones can be explained by the purpose behind the design of these webpages. Commercial webpages try to write their information using simple texts to achieve

higher sales of hearing aids and other commercial benefits. However, governmental webpages write more about funding available to the population and the candidacy to receive hearing aids or cochlear implants. The findings of my study therefore support the second hypothesis.

In my third hypothesis, it was hypothesised that a significant number of webpages are not HON certified. The results of my study support this hypothesis because there was no webpage which was HON certified. Similar findings are reported by Hsu (2017). Although there is no cost to become HON certified (Health On the Net Foundation, 2010), none of the Hindi webpages are HON certified. The factors contributing to this finding may be the awareness and the importance provided to this subject. Research has confirmed that health webpages with HON certification provide better quality information than the ones without HON certification (Breckons, Jones, Morris, & Richardson, 2008; Nason, Byrne, Noel, Moore, & Kiely, 2012). Therefore, it can be suggested that the developers and owners of Hindi hearing-related webpages do not care about the quality of their webpage. Another reason might be that they are not aware of this tool.

It was hypothesised that there is a significant difference in readability score obtained by RH1 and RH2 readability formulas. The results supported this hypothesis because readability scores obtained by RH1 were significantly higher than those obtained by RH2. There is no literature to support or contradict these findings. But a possible explanation could be in the structure of their formulas: RH1 ($-2.34 + 2.14 * AWL + 0.01 PSW$) and RH2 ($0.211 + 1.37 * AWL + 0.005 * JUK$) were designed by Sinha et al. (2012) due to the inapplicability of English formulas on Hindi text (Agnihotri & Khanna, 1991; Bhagoliwal, 1961). Sinha et al. (2012) observed that AWL, JUK, PSW and PSW30 are key features contributing towards readability in Hindi.

This study's next hypothesis was to find if there is a significant correlation between the readability scores calculated by RH1 and RH2. My study results indicated a positive significant

correlation between the readability scores of the RH1 and RH2 formulas. Due to unavailability of research on the relationship between these two readability formulas, I cannot comment on the consistency or contradiction of my findings. However, the results explain that either of these formulas can be used to calculate the readability score of a Hindi text, due to the strong, positive correlation between them, supporting the hypothesis.

4.2.2 Testing readability ease by a cloze test

Although the readability scores obtained by using software and the readability formulas provide a number to predict the readability ease of a document, these numbers cannot predict the understandability of the paragraph or a text (Atcherson, Zraick, & Brasseux, 2011; Kahn & Pannbacker, 2000). Analysing the readability by focussing on the understanding of the text illustrate the uniqueness of this research, as, in the past, readability research reviewers have been unable to find good readability studies based on these constructs (Horner, Surratt, & Juliusson, 2000). Therefore, in my study, I used a measure called a cloze test to find more information regarding readability ease of Hindi text based on the comprehensibility of the text.

The statistical analysis did not support Hypothesis 6, revealing no significant difference between the cloze scores and RGLs of the webpages. The mean cloze score of the text with maximum mean RGL and with minimum mean RGL (both obtained by readability formulas) were 50% and 43.3%, respectively, which suggests that the online hearing material available in Hindi text can be used, but may require some additional guidance or teaching (Doak, 1996). It suggests that the paragraph with highest mean RGL is as understandable as the paragraph with lowest RGL, which is not a finding I expected. Moreover, if we look at the effect size ($d = 0.40$), it shows a difference between the findings of two paragraphs. But the lack of statistical significance could be due to the small sample size and the absence of normal distribution of the data.

Another possible reason behind these findings is that the validated readability formulas need further evaluation, and so it is necessary to identify if they can be applied to a hearing-related text in Hindi. It may be possible that the participants of my study are not representative of consumers of online hearing information in Hindi because they were well educated and fairly young. No research was found in the literature on the question of readability of online hearing-related information in Hindi obtained by a cloze test.

The cloze scores of paragraphs with maximum readability showed SD values of 11.36 and with minimum readability showed SD values of 5.06, which are likely to be related to health literacy. Although the participants of this study were young and well educated, health literacy cannot be judged by these factors. According to some studies (Joubert & Githinji, 2014; Kahn & Pannbacker, 2000), an unseen problem of low health literacy has an effect on people with different personalities and cannot be estimated by looks, level of education (i.e. years of school education) or their economic background. Moreover, literature has shown that literacy skills can change, depending on the familiarity of the written text material (Joubert & Githinji, 2014). It suggests that if a person knows the context of the written paragraph, the person may show excellent literacy skills. To address this issue, in my study, all the participants did not have a background in hearing science, because somebody from a science background could understand the paragraph more easily and could show better literacy skills. As far as the readability of online hearing material is concerned, it is important to know the audience and their literacy skills. For an online hearing-related information provider, it is difficult to know the context in which readers struggle the most.

Another important factor influencing the cloze scores can be lack of attention. It is an important factor because it initiates the chemical reactions activating nerve impulses to send messages to the brain (Doak, 1996). If participants do not pay attention to the paragraph while filling the blanks in the cloze test, results can vary. In my study, I observed that participants

tried to finish the task and did not pay attention to the text. Therefore, it could be a possible reason to have a low percentage of correct words.

According to Doak (1996), if the logic, language, and experience built into any health-related information does not match with the logic, language and experience of the reader, understanding of the information is greatly impacted. For example, if there is no logic behind the use of hearing aids, it is highly possible that readers will stop reading the information given. Language mismatching means the use of unfamiliar words. It is commonly seen that in hearing information, words such as cochlea, vestibular system, tympanic membrane are used. It can impact the ease of readability. In addition, readability also reduces when familiar words are used in unfamiliar contexts. Language played a major role in this study, because in India English is the more common language used by youngsters to search for any kind of health-related topic.

4.3 Clinical implications

As it is becoming common that people search the Internet when they require any hearing-related information in India, especially in Hindi (Akerkar et al., 2005), this study was conducted to assess the readability ease of the hearing-related information available on the Internet in Hindi. The findings of this study suggest that the readability for the hearing-related webpages in Hindi cannot be assessed by readability formulas exclusively because they do not consider reader's literacy skills and context of the text (Redish, 2000). This is important when clinicians provide any website or any information that is available online to the consumer.

Prior studies have noted the importance of superior health material provided by health care personnel or organisations in achieving consumer's satisfaction (Bylund et al., 2007; McMullan, 2006). Furthermore, superior quality materials increase the self-confidence of clients (Sommerhalder, Abraham, Zufferey, Barth, & Abel, 2009). High quality can be achieved by making sure that the material is easy to read and understandable, the information

accessed by the client is relevant, and assess the need of rehabilitation (Laplante-Lévesque et al., 2012). Clinicians should be careful when recommending any Internet information to a client and should be aware of its quality and readability ease (Laplante-Lévesque et al., 2012; Laplante-Lévesque & Thorén, 2015). HON certification and DISCERN can be used to assess the quality and suitability of the website (Charnock, Shepperd, Needham, & Gann, 1999; Health On the Net Foundation, 2010). This study's findings demonstrate nonavailability of HON certified webpages in Hindi, which means that the quality of online hearing-related webpages in Hindi cannot be confirmed. Therefore, clinicians should be careful before recommending any online hearing-related information to their clients.

4.4 Study limitations

Despite having interesting results from this study, there are some limitations which need to be discussed. In the following subsections, the limitations of each step of my study will be discussed.

4.4.1 Readability search and webpages in Hindi

In this study, efforts were made to replicate the search strategy that had been used in other readability studies performed in different languages (Hsu, 2017) so that results could be compared. However, it is possible that different search key terms in a different style can be used by people speaking different languages. This can further impact the results of the webpages obtained by the search using specific key terms. Another limitation in the search strategy was that the search terms were decided upon by putting up a question on the Facebook friend list, because it could represent people belonging to a particular demographic group. In addition, it is not necessary that each person in the target population uses Facebook. In India, only 15% of the whole population use Facebook (Statista, 2017), therefore, the people recruited by Facebook cannot represent the whole Indian population.

Another limitation of my study is the low number of hearing webpages available in Hindi compared to those available in English. Around 80% of Web information was available in English until 1990, and by 2011, the information extended to Chinese, French, German, Russian and Spanish ,but Indian languages still lagged behind (Bhattacharya, 2017). Online hearing-related information in Hindi became available recently, which can explain the low number of online Hindi hearing-related webpages. In addition, when I searched for webpages in Hindi, most of the webpages used English terminology but written in Hindi alphabets, and two webpages were completely written in Hindi alphabets with English terminology – these were excluded from the study. Selecting the content in Hindi caused difficulties in identifying the readability of that webpage; therefore, the sample size of my study was small and I used a cloze test to get more information about the readability ease of Hindi text related to hearing available on the Internet.

4.4.2 Readability Formulas

Many researchers in the past have talked about the complete trust in readability formulas for identifying the readability ease of a text. In a literature review conducted by Redish (2000), authors observed numerous technical limitations of readability formulas and suggested not relying completely on them to find the readability ease of any type of text. They also mentioned that readability formulas were developed for text books for children, not for technical text. They ignored the importance of reader literacy and text comprehension which I discussed earlier in section 4.2.1. In another study, performed by Bailin and Grafstein (2001), the authors argued that “readability formulas do not measure in fact what they are designed to measure” (p. 298). The possible explanation behind this is that readability formulas are based on linguistic assumptions. Although these studies presented arguments for English readability formulas, this can be applicable to the formulas in Hindi. As already explained in the Introduction section, Hindi readability formulas are based on textual features such as average

word length, number of syllables in a word and consonant adjuncts. They ignore the word meaning, its context and the reader's literacy skills. In this study, RH1 and RH2 formulas (Sinha et al., 2012) were used to identify the readability of Hindi text because these formulas were designed by using support vector models and support vector regressions (Sinha, Dasgupta, & Basu, 2014). They also included six syntactic and lexical features in these formulas (Sinha et al., 2014,). But these formulas did not consider the difficulty of health-related information, because some health-related webpages use English terminology to explain the issues. In addition, readability formulas do not include the method of presentation of the information. Some people feel comfortable with reading the information on a pamphlet but not on a computer (Meyer & Hickson, 2012).

Another limitation of the RH1 and RH2 formulas was the computational model. It was not very easy to operate and could only give the textual features by parsing. Readability score was calculated later manually. Moreover, for a few web pages, the parser did not work and parsing was also done manually, which increased the chances of error and bias.

4.4.3 Cloze test

In this study, the limitation was the recruitment of participants for the cloze test. All the participants were young and from a high socioeconomic background which increased the bias because, in reality, the target population is older and from a lower socioeconomic status. In addition, the participants struggled with the Hindi language because they always use English language to read webpages. Also, they did not pay full attention to the paragraph and tried to finish the task in a couple of minutes which also impacted the results. Although no participant was recruited from a hearing background, a few were from a science background and some were from arts and commerce backgrounds.

4.4.4 Other limitations

Due to the unavailability of readability research in Hindi, it was difficult to compare the results of this study with others. There is no literature available that discusses the validity of the readability formulas in Hindi. Sinha et al. (2012) designed these formulas and validated them. No other researcher has validated these formulas. But researchers in different user groups have mentioned that they are not validated.

4.5 Future research

4.5.1 New models for readability formulas

This study is the first step towards the readability analysis of hearing-related information available on the Internet in Hindi. Because the readability formulas in Hindi were not designed for health-related webpages, this study was not able to practically assess the readability grade level of hearing-related webpages. Therefore, future research is required to update the present readability formulas so that health information can also be assessed at the level of reading difficulty. The Google search engine in Hindi is becoming popular among the Indian population and webpages containing health-related content in Hindi are also increasing day by day (Statista, 2017), which clearly indicates the need for more research.

4.5.2 Health On the Net (HON) Code and DISCERN

In this study, none of the webpages were identified as HON certified. As has been explained in previous sections, HON code certification requires the completion of some application forms and the adoption of the principles of HON certification (Health On the Net Foundation, 2010). Even though the certification is free, no webpage owners had made an effort to gain the certification. It could possibly be due to lack of awareness of this code rather than indicating a poor quality webpage. It has been observed in the previous literature that the webpages with HON certification showed superior quality than those without HON certification (Breckons et al., 2008; Nason et al., 2012). Now, there is an automated instrument

available which is designed by the HON foundation to make this certification quick and automatic (Boyer et al., 2015). However, it cannot be used for webpages in Hindi, which could possible cause the lower adoption rate. This is another step for future research and development in the field of readability analysis of Hindi online hearing-related webpages that can improve the quality of the information provided on webpages.

DISCERN is another tool comprising 16 elements which also helps in the assessment of the quality of health information (Charnock et al., 1999). The literature has confirmed its reliability (Rees, Ford, & Sheard, 2002) and its use to assess the quality of health information (Laplante-Lévesque et al., 2012; Ritchie, Tornari, Patel, & Lakhani, 2016). This tool has not been designed for languages other than English (Charnock et al., 1999). This could be another step for future research – to develop a similar tool for Hindi readability analysis.

4.5.3 Focus on the readability of important sections of the webpage

It is possible that some parts of a webpage are easy to read and that the crucial part about health information, such as the aetiology of a disease, its signs and symptoms may be harder to read. These sections are very important for consumers seeking health information and should be easy to understand by the reader. Researchers have suggested focussing on the readability analysis of these important sections so that it can help consumers to understand the information without any misunderstanding (Wang, Miller, Schmitt, & Wen, 2013). In this study, random content was selected from the webpage for the readability analysis and no important parts were selected to analyse their readability. The focus on making the important sections of the webpage with low RGL can help to increase the understanding of health information in the future.

4.5.4 Focus on readability and comprehension

In this study, although a cloze test and readability formulas were used to assess the readability ease of online hearing-related text, there are other factors that influence the

readability which were not included in this study (Friedman & Hoffman-Goetz, 2006; Klare, 1976) such as writing style, comprehension and health literacy. To date, there are fewer studies available in the readability literature where readability is assessed by both formulas and a comprehension method. Therefore, this is an important area for future research that could increase our knowledge about the reading ease of the health material available on Internet.

4.6 Conclusions

This study identified the readability ease of online hearing-related information available in Hindi available to consumers who speak Hindi as their first language. Readability was analysed using the RH1 and RH2 readability formulas proposed by Sinha et al. (2012) and a cloze test.

The results of the study demonstrated that RGL calculated by the formulas was within the recommended value, which means the hearing-related material available on the Internet in Hindi is easy to read. However, the results of readability ease calculated by the cloze test suggested that the paragraphs with maximum RGL and minimum RGL were not significantly different from each other in their level of difficulty in understanding.

From a clinical perspective, it means that clinicians should be careful before recommending any online hearing material to their patients based on the RGL. Moreover, readability formulas should be evaluated further for the specific user population and the content of the information provided on Internet. Also, quality and suitability of the website should also be determined by the web developers by using various tools such as HON and DISCERN in order to increase the trust of consumers in the information provided by the website designers.

5 Reference List

- Agnihotri, R. K., & Khanna, A. L. (1991). Evaluating the readability of school textbooks: An Indian study. *Journal of Reading*, 35(4), 282-288.
- Akerkar, S. M., Kanitkar, M., & Bichile, L. (2005). Use of the Internet as a resource of health information by patients: A clinic-based study in the Indian population. *Journal of Postgraduate Medicine*, 51(2), 116.
- American Speech-Language-Hearing Association. (2015). Type, degree, and configuration of hearing loss. *Audiology Information Series*. Retrieved from <http://www.asha.org/uploadedFiles/AIS-Hearing-Loss-Types-Degree-Configuration.pdf>
- American Speech-Language-Hearing Association ASHA. (1981). On the definition of hearing handicap.
- Appollonio, I., Carabellese, C., Frattola, L., & Trabucchi, M. (1996). Effects of sensory aids on the quality of life and mortality of elderly people: A multivariate analysis. *Age and Ageing*, 25(2), 89-96.
- Atcherson, S. R., DeLaune, A. E., Hadden, K., Zraick, R. I., Kelly-Campbell, R. J., & Minaya, C. P. (2014). A Computer-based Readability Analysis of Consumer Materials on the American Speech-Language-Hearing Association Website. *Contemporary Issues in Communication Science & Disorders*, 41.
- Atcherson, S. R., Zraick, R. I., & Brasseux, R. E. (2011). Readability of patient-reported outcome questionnaires for use with persons with tinnitus. *Ear and Hearing*, 32(5), 671-673.
- Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language and Communication*, 21(3), 285-301. doi: 10.1016/S0271-5309(01)00005-2

- Baujard, V., Boyer, C., & Geissbühler, A. (2010). Evolution of health web certification, through the HONcode experience. *Swiss Medical Informatics*, 26(69), 53-55.
- Begg, E. J., Barclay, M. L., & Kirkpatrick, C. M. (2001). The therapeutic monitoring of antimicrobial agents. *British Journal of Clinical Pharmacology*, 52(S1), 35-43.
- Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: An updated systematic review. *Annals of Internal Medicine*, 155(2), 97-107. doi: 10.7326/0003-4819-155-2-201107190-00005
- Bernstam, E. V., Shelton, D. M., Walji, M., & Meric-Bernstam, F. (2005). Instruments to assess the quality of health information on the World Wide Web: What can our patients actually use? *International Journal of Medical Informatics*, 74(1), 13-19. doi: 10.1016/j.ijmedinf.2004.10.001
- Bhagoliwal, B. (1961). Readability formulae: Their reliability, validity and applicability in hindi. *Journal of Education and Psychology*, 19, 13-26.
- Bhattacharya, A. (2017). India's internet users have more faith in content that's not in English. Retrieved from <https://qz.com/972844/indias-internet-users-have-more-faith-in-content-thats-not-in-english-study-says/>
- Black, H. C. (1910). The Law Dictionary , featuring Black's Law Dictionary free online legal dictionary. <http://thelawdictionary.org/reading-grade-level-rgl/>
- Bolz, H. J. (2016). Hereditary hearing loss and Its Syndromes *European Journal of Human Genetics*, 24(11), 1650. doi: 10.1038/ejhg.2016.67
- Boothroyd, A. (2007). Adult Aural Rehabilitation: What is it and does it work? *Trends in Amplification*, 11(2), 63-71. doi: 10.1177/1084713807301073
- Boyer, C., Dolamic, L., Ranasinghe, M., & Baujard, V. (2015). An automated HONcode detection system informs internet users of HONcode compliance. *Swiss Medical Informatics*, 31.

- Breckons, M., Jones, R., Morris, J., & Richardson, J. (2008). What do evaluation instruments tell us about the quality of complementary medicine information on the internet? *Journal of Medical Internet Research*, 10(1).
- Bylund, C. L., Gueguen, J. A., Sabee, C. M., Imes, R. S., Li, Y., & Sanford, A. A. (2007). Provider–patient dialogue about internet health information: An exploration of strategies to improve the provider–patient relationship. *Patient Education and Counseling*, 66(3), 346-352.
- Carlbring, P., & Andersson, G. (2006). Internet and psychological treatment. How well can they be combined? *Computers in Human Behavior*, 22(3), 545-553.
- Carter, B. (2000). Formula for failure: Reading levels and readability formulas do not create lifelong readers. *School Library Journal*, 46(7), 34-37.
- Charnock, D., Shepperd, S., Needham, G., & Gann, R. (1999). DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *Journal of Epidemiology and Community Health*, 53(2), 105-111.
- Chinn, M. D., & Fairlie, R. W. (2007). The determinants of the global digital divide: A cross-country analysis of computer and internet penetration. *Oxford Economic Papers*, 59(1), 16-44. doi: 10.1093/oep/gpl024
- Chisolm, T. H., Johnson, C. E., Danhauer, J. L., Portz, L. J., Abrams, H. B., Lesner, S., . . . Newman, C. W. (2007). A systematic review of health-related quality of life and hearing aids: Final report of the American Academy of Audiology Task Force On the Health-Related Quality of Life Benefits of Amplification in Adults. *Journal of the American Academy of Audiology*, 18(2), 151-183.
- Clausing, G., & Senko, D. (1978). On the scoring of cloze tests. *Die Unterrichtspraxis / Teaching German*, 11(2), 74-81. doi: 10.2307/3529214

- Cline, R. J. W., & Haynes, K. M. (2001). Consumer health information seeking on the internet: The state of the art. *Health Education Research*, 16(6), 671-692. doi: 10.1093/her/16.6.671
- Cox, R. M., & Alexander, G. C. (2002). The International Outcome Inventory for Hearing Aids (IOI-HA): Psychometric properties of the English version: El Inventario International de Resultados para Auxiliares Auditivos (IOI-HA): Propiedades psicometricas de la version en ingles. *International Journal of Audiology*, 41(1), 30-35.
- Dale, E., & Chall, J. S. (1949). The concept of readability. *Elementary English*, 26(1), 19-26.
- Dalton, D. S., Cruickshanks, K. J., Klein, B. E., Klein, R., Wiley, T. L., & Nondahl, D. M. (2003). The impact of hearing loss on quality of life in older adults. *The Gerontologist*, 43(5), 661-668.
- Davison, A. (1999). *Syntax and morphology in Hindi and Urdu: A lexical resource*. Retrieved from <https://clas.uiowa.edu/linguistics/hindi-verb-project>
- DeWalt, D. A., Berkman, N. D., Sheridan, S., Lohr, K. N., & Pignone, M. P. (2004). Literacy and health outcomes: A systematic review of the literature. *Journal of General Internal Medicine*, 19(12), 1228-1239. doi: 10.1111/j.1525-1497.2004.40153.x
- Diaz, J. A., Griffith, R. A., Ng, J. J., Reinert, S. E., Friedmann, P. D., & Moulton, A. W. (2002). Patients' use of the internet for medical information. *Journal of General Internal Medicine*, 17(3), 180-185.
- Doak, C. C., Doak L.G., Root J. H. (1996). *Teaching patients with low literacy skills* (2nd ed.). Philadelphia,PA: J. B. Lippincott Company.
- DuBay, W. H. (2004). *The Principles of Readability*. Costa Mesa, CA: Impact Information.

- Eriksson-Mangold, M., & Carlsson, S. (1991). Psychological and somatic distress in relation to perceived hearing disability, hearing handicap, and hearing measurements. *Journal of Psychosomatic Research*, 35(6), 729-740.
- Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ*, 324(7337), 573-577.
- Eysenbach, G., Powell, J., Kuss, O., & Sa, E.-R. (2002). Empirical studies assessing the quality of health information for consumers on the world wide web: A systematic review. *JAMA*, 287(20), 2691-2700.
- Fiksdal, A. S., Kumbamu, A., Jadhav, A. S., Cocos, C., Nelsen, L. A., Pathak, J., & McCormick, J. B. (2014). Evaluating the process of online health information searching: A qualitative approach to exploring consumer perspectives. *Journal of Medical Internet Research*, 16(10), e224. doi: 10.2196/jmir.3341
- Finitzo, T., & Crumley, W. G. (1999). The role of the pediatrician in hearing loss: From detection to connection. *Pediatric Clinics of North America*, 46(1), 15-34.
- Fox, S. (2012). Online health search 2006. Pew Internet and American Life Project. 2006 Oct 29. Retrieved from <http://www.pewinternet.org/2006/10/29/online-health-search-2006/>
- Friedman, D. B., & Hoffman-Goetz, L. (2006). A systematic review of readability and comprehension instruments used for print and web-based cancer information. *Health Education and Behavior*, 33(3), 352-373.
- Garg, S., Chadha, S., Malhotra, S., & Agarwal, A. (2009). Deafness: Burden, prevention and control in India.

Google.org. (2016). Google 'Health Speaks' for online medical info in Hindi, THE HINDU.

Retrieved from <http://www.thehindu.com/sci-tech/technology/Google-Health-Speaks-for-online-medical-info-in-Hindi/article15906333.ece>

Gordon-Salant, S., Frisina, R. D., Fay, R. R., & Popper, A. (2010). The Aging Auditory System (Vol. 34). New York, NY: Springer Science & Business Media.

Health On the Net Foundation, H. (2010). Operational definition of the HONcode principles

Health On the Net Code of Conduct Retrieved from

<http://www.healthonnet.org/HONcode/Guidelines/guidelines.html>

Hickson, L., Worrall, L., & Scarinci, N. (2007). A randomized controlled trial evaluating the active communication education program for older people with hearing impairment.

Ear and Hearing, 28(2), 212-230.

Horner, S. D., Surratt, D., & Juliusson, S. (2000). Improving readability of patient education materials. Journal of Community Health Nursing, 17(1), 15-23.

Hsu, P.-H. (2017). Readability of hearing related internet information in Traditional Chinese.

University of Canterbury, Christchurch.

IBM. (2016). IBM SPSS Statistics (Version 24). United States of America. Retrieved from

<https://www.ibm.com/products/spss-statistics>

India, T. R. C. o. (1992). The Rehabilitation Council of India Act. Retrieved 11 May 2017

<http://www.rehabcouncil.nic.in/engweb/rciact.pdf>

Internet World Stats. (2015). Alphabetical List of Countries. Retrieved from

<http://www.internetworldstats.com/list2.htm>

Internet World Stats. (2017a). Surfing and Site Guide. Retrieved from

<http://www.internetworldstats.com/surfing.htm> - 1

Internet World Stats. (2017b). Usage and population statistics. Retrieved from

<http://www.internetworldstats.com/top25.htm>

- Internet World Stats. (2017c). World Internet users and 2017 population stats. Retrieved from <http://www.internetworldstats.com/>
- Joubert, K., & Githinji, E. (2014). Quality and readability of information pamphlets on hearing and paediatric hearing loss in the Gauteng Province, South Africa. *International Journal of Pediatric Otorhinolaryngology*, 78(2), 354-358.
- Kachru, Y. (2006). Hindi (Vol. 12). Philadelphia, PA: John Benjamins.
- Kahn, A., & Pannbacker, M. (2000). Readability of educational materials for clients with cleft lip/palate and their families. *American Journal of Speech-Language Pathology*, 9(1), 3-9.
- Kellogg, S. H. (1876). A grammar of Hindi Language. Allahabad, AD: Am. Pres. Mission Press.
- Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. from DTIC Document <http://www.dtic.mil/dtic/tr/fulltext/u2/a006655.pdf>
- Klare, G. R. (1976). A second look at the validity of readability formulas. *Journal of reading behavior*, 8(2), 129-152.
- Kochkin, S., & Rogin, C. M. (2000). Quantifying the obvious: The impact of hearing instruments on quality of life. *Hearing Review*, 7(1), 6-34.
- Laforge, R. G., Spector, W. D., & Sternberg, J. (1992). The relationship of vision and hearing impairment to one-year mortality and functional decline. *Journal of Aging and Health*, 4(1), 126-148.
- Laplane-Lévesque, A., Kathleen Pichora-Fuller, M., & Gagné, J.-P. (2006). Providing an internet-based audiological counselling programme to new hearing aid users: A qualitative study: Provisión de un programa de consejería audiológica por Internet

- para nuevos usuarios de auxiliares auditivos: un estudio cualitativo. *International Journal of Audiology*, 45(12), 697-706.
- Laplante-Lévesque, A., Knudsen, L. V., Preminger, J. E., Jones, L., Nielsen, C., Öberg, M., . . . Kramer, S. E. (2012). Hearing help-seeking and rehabilitation: Perspectives of adults with hearing impairment. *International Journal of Audiology*, 51(2), 93-102.
- Laplante-Lévesque, A., & Thorén, E. S. (2015). Readability of Internet information on hearing: Systematic literature review. *AMERICAN JOURNAL OF AUDIOLOGY*, 24(3), 284-288.
- Lewis-Cullinan, C., & Janken, J. K. (1990). Effect of cerumen removal on the hearing ability of geriatric patients. *Journal of Advanced Nursing*, 15(5), 594-600.
- Madell, J. R. (2014a). *Pediatric Audiology : Diagnosis, Technology and Management*. In C. F. Jane R. Madell (Ed.), *Pediatric* (Second ed., pp. 89-94). New York, NY 10001: Thieme Medical Publishers, Inc.
- Madell, J. R. (2014b). *Pediatric audiology: Diagnosis, technology, and management*. In J. R. F. C. Madell (Ed.), *Pediatric Audiology: Diagnosis, Technology, and Management* (2nd ed., pp. 79-88). New York, NY Thieme Medical Publishers.
- Marques, A., Kozlowski, L., & Marques, J. M. (2004). Reabilitação auditiva no idoso. *Revista Brasileira de Otorrinolaringologia*, 70, 806-811.
- Martin, F. N., & Clark, J. G. (2012). *Introduction to audiology*. Philadelphia, NJ: Pearson Education.
- McLaughlin, G. H. (1969). SMOG grading-a new readability formula. *Journal of Reading*, 12(8), 639-646.
- McMullan, M. (2006). Patients using the internet to obtain health information: How this affects the patient–health professional relationship. *Patient Education and Counseling*, 63(1), 24-28.

- Mendel, L. L. (2007). Objective and subjective hearing aid assessment outcomes. AMERICAN JOURNAL OF AUDIOLOGY, 16(2), 118-129. doi: 10.1044/1059-0889(2007/016)
- Meyer, C., & Hickson, L. (2012). What factors influence help-seeking for hearing impairment and hearing aid adoption in older adults? International Journal of Audiology, 51(2), 66-74.
- Ministry of Home Affairs, G. o. I. (2001). Literacy and Level of Education. Retrieved from http://www.censusindia.gov.in/Census_And_You/literacy_and_level_of_education.aspx
- Ministry of Home Affairs, G. o. I. (2011). Provisional Population Totals : India : Census 2011. Retrieved from <http://www.censusindia.gov.in/2011-prov-results/indiaatglance.html>
- Møller, A. R. (2012). Hearing: anatomy, physiology, and disorders of the auditory system: Plural Publishing.
- Montano, J. J., Montano, J., & Spitzer, J. (2009). Defining audiologic rehabilitation. Adult audiologic rehabilitation, 26-36.
- Moreland, J., French, T. L., & Cumming, G. P. (2015). The prevalence of online health information seeking among patients in scotland: a cross-sectional exploratory study. JMIR research protocols, 4(3), e85.
- Moss, W. J., Lin, H. W., & Cueva, R. A. (2016). Surgical and audiometric outcomes for repair of congenital aural atresia and hypoplasia. JAMA Otolaryngology–Head & Neck Surgery, 142(1), 52-57.
- Mulrow, C. D., Aguilar, C., Endicott, J. E., Velez, R., Tuley, M. R., Charlip, W. S., & Hill, J. A. (1990). Association between hearing impairment and the quality of life of elderly individuals. Journal of the American Geriatrics Society, 38(1), 45-50.

- Musiek, F. E., & Baran, J. A. (1986). Neuroanatomy, neurophysiology, and central auditory assessment. Part I: Brain stem. *Ear and Hearing*, 7(4), 207-219.
- Nason, G. J., Byrne, D. P., Noel, J., Moore, D., & Kiely, P. J. (2012). Scoliosis-specific information on the internet: Has the “information highway” led to better information provision? *Spine*, 37(21), E1364-E1369.
- National Center for Education Statistics. (2003). National assessment of adult literacy. Retrieved from <https://nces.ed.gov/naal/>
- National Sample Survey Organization. (2003). Disability persons in India . NSS 58th round (July-December 2002). Retrieved from file:///Users/mac/Downloads/ddi-documentation-english-196.pdf
- Newton, V. E., & Shah, S. R. (2013). Improving communication with patients with a hearing impairment. *Community Eye Health*, 26(81), 6-7.
- Öberg, M., Wänström, G., Hjertman, H., Lunner, T., & Andersson, G. (2009). Development and initial validation of the ‘Clinical Global Impression’ to measure outcomes for audiological rehabilitation. *Disability and Rehabilitation*, 31(17), 1409-1417.
- Office of the Registrar General & Census Commissioner. (2001). Disabled population. Retrieved from http://censusindia.gov.in/Census_And_You/disabled_population.aspx
- ORACLE. (2015). The Java™ tutorials. Retrieved from <http://docs.oracle.com/javase/tutorial/uiswing/>
- Pacala, J. T., & Yueh, B. (2012). Hearing deficits in the older patient: “I didn't notice anything”. *JAMA*, 307(11), 1185-1194.
- Popelka, M. M., Cruickshanks, K. J., Wiley, T. L., Tweed, T. S., Klein, B. E., & Klein, R. (1998). Low prevalence of hearing aid use among older adults with hearing loss: The epidemiology of hearing loss study. *Journal of the American Geriatrics Society*, 46(9), 1075-1078.

- Press Trust of India. (2017, April 30, 2017). 536 million users to log onto internet in Indian languages by 2021, The Economic Times. Retrieved from <http://economictimes.indiatimes.com/tech/internet/536-mn-users-to-log-onto-internet-in-indian-languages-by-2021/articleshow/58442659.cms>
- Redish, J. (2000). Readability formulas have even more limitations than Klare discusses. *ACM Journal of Computer Documentation (JCD)*, 24(3), 132-137.
- Rees, C. E., Ford, J. E., & Sheard, C. E. (2002). Evaluating the reliability of DISCERN: A tool for assessing the quality of written patient information on treatment choices. *Patient Education and Counseling*, 47(3), 273-275.
- Rehab Council of India. (1995). Hearing impairment. New Delhi, ND: Retrieved from <http://www.rehabcouncil.nic.in/writereaddata/hi.pdf>.
- Ritchie, L., Tornari, C., Patel, P., & Lakhani, R. (2016). Glue ear: How good is the information on the World Wide Web? *The Journal of Laryngology and Otology*, 130(2), 157.
- Samuel, S. E., & Babu, L. G. (2017). Association of severity of auditory impairment related handicap with quality of life in older adults. *Indian Journal of Physiotherapy and Occupational Therapy*, 11(1), 80-84.
- Santos, R. B. F., Marangoni, A. T., de Andrade, A. N., Prestes, R., & Gil, D. (2014). Effects of auditory training in individuals with high-frequency hearing loss. *Clinics*, 69(12), 835-840. doi: 10.6061/clinics/2014(12)08
- Schlauch, R. S., & Nelson, P. (2015). Puretone evaluation. In J. Katz , Chasin , M., English , K., Hood ,L. J., & Tillery, K. L. (Ed.), *Handbook of clinical audiology* (7th Edition ed., pp. 29-47). Philadelphia, PA: Wolters Kluwer.
- Shahnazari-Dorcheh, M., Roshan, S., & Hesabi, A. (2012). What Is the optimum length of a cloze test? *International Journal of English Linguistics*, 2(5), 142.

- Sharma, S. (2012). Studying readability of hindi texts: A computational approach. Master of Technology, Indian Institute of Technology, Kharagpur.
- Sinha, M., Dasgupta, T., & Basu, A. (2014). Text Readability in Hindi: A Comparative Study of Feature Performances Using Support Vectors. Paper presented at the 11th International Conference on Natural Language Processing.
- Sinha, M., Dasgupta, T., & Basu, A. (2014,). Design and Development of an Online Computational Framework to Facilitate Language Comprehension Research on Indian Languages.
- Sinha, M., Sharma, S., Dasgupta, T., & Basu, A. (2012). New readability measures for Bangla and Hindi texts. <http://www.aclweb.org/anthology/C12-2111>
- Sommerhalder, K., Abraham, A., Zufferey, M. C., Barth, J., & Abel, T. (2009). Internet information and medical consultations: Experiences from patients' and physicians' perspectives. *Patient Education and Counseling*, 77(2), 266-271.
- Sørensen, K., Van den Broucke, S., Fullam, J., Doyle, G., Pelikan, J., Slonska, Z., & Brand, H. (2012). Health literacy and public health: A systematic review and integration of definitions and models. *BMC Public Health*, 12(1), 80.
- Statista. (2017). Number of social network users in India from 2015 to 2022 (in millions). Retrieved from <https://www.statista.com/statistics/278407/number-of-social-network-users-in-india/>
- Steiger, J. R. (2015). Handbook of clinical audiology. In J. Katz (Ed.), (7th edition ed., pp. 49-59). Philadelphia, PA: Wolters Kluwer.
- Stephens, D., Lewis, P., Davis, A., Gianopoulos, I., & Vetter, N. (2001). Hearing aid possession in the population: Lessons from a small country: Posesión de auxiliares auditivos en la población: Lecciones de un pequeño país. *Audiology*, 40(2), 104-111.

- Stevens, G., Flaxman, S., Brunskill, E., Mascarenhas, M., Mathers, C. D., & Finucane, M. (2013). Global and regional hearing impairment prevalence: An analysis of 42 studies in 29 countries. *The European Journal of Public Health*, 23(1), 146-152.
- Stokroos, R., Albers, F., & Tenvergert, E. (1998). Antiviral treatment of idiopathic sudden sensorineural hearing loss: A prospective, randomized, double-blind clinical trial. *Acta Oto-Laryngologica*, 118(4), 488-495.
- Takahashi, G., Martinez, C. D., Beamer, S., Bridges, J., Noffsinger, D., Sugiura, K., . . . Williams, D. W. (2007). Subjective measures of hearing aid benefit and satisfaction in the NIDCD/VA follow-up study. *Journal of the American Academy of Audiology*, 18(4), 323-349.
- Taylor, W. L. (1953). "Cloze procedure": a new tool for measuring readability. *Journalism Bulletin*, 30(4), 415-433.
- Tonsaker, T., Bartlett, G., & Trpkov, C. (2014). Health information on the Internet Gold mine or minefield? *Canadian Family Physician*, 60(5), 407-408.
- UCLA Language. (2014). Hindi. Retrieved from <http://www.lmp.ucla.edu/Profile.aspx?menu=004&LangID=87>
- Union, I. T. (2016, 15 September 2016). Committed to connecting the world. Retrieved from <http://www.itu.int/en/mediacentre/Pages/2016-PR35.aspx>
- Union, I. T. (2016, 15 September 2016). The State of Broadband 2016: Broadband Catalyzing Sustainable Development. Retrieved from https://www.itu.int/dms_pub/itu-s/opb/pol/S-POL-BROADBAND.17-2016-PDF-E.pdf
- van der Heide, I., Wang, J., Droomers, M., Spreeuwenberg, P., Rademakers, J., & Uiters, E. (2013). The relationship between health, education, and health literacy: results from

- the Dutch Adult Literacy and Life Skills Survey. *Journal of health communication*, 18(sup1), 172-184.
- Varshney, S. (2016). Deafness in India. *Indian Journal of Otology*, 22(2), 73.
- Vestergaard Knudsen, L., Öberg, M., Nielsen, C., Naylor, G., & Kramer, S. E. (2010). Factors influencing help seeking, hearing aid uptake, hearing aid use and satisfaction with hearing aids: A review of the literature. *Trends in Amplification*, 14(3), 127-154.
- Vikram, S. (2013). Morphology: Indian Languages and European Languages. *International Journal of Scientific and Research Publications*, 3(6).
- Vishwambhar, S. (2015). Hearing in India: All aspects. *Online Journal of Otolaryngology*, 5(1), 39.
- Walsh, T. M., & Volsko, T. A. (2008). Readability assessment of internet-based consumer health information. *Respiratory Care*, 53(10), 1310-1315.
- Wang, L.-W., Miller, M. J., Schmitt, M. R., & Wen, F. K. (2013). Assessing readability formula differences with written health information materials: application, results, and recommendations. *Research in Social and Administrative Pharmacy*, 9(5), 503-516.
- Weiss, B. D., & Natl Work Grp Literacy, H. (1998). Communicating with patients who have limited literacy skills: Report of the national work group on literacy and health. *Journal of Family Practice*, 46(2), 168-176.
- Williams, R. L., Chalmers, T. C., Stange, K. C., Chalmers, F. T., & Bowlin, S. J. (1993). Use of antibiotics in preventing recurrent acute otitis media and in treating otitis media with effusion: A meta-analytic attempt to resolve the brouhaha. *JAMA*, 270(11), 1344-1351.
- World Health Organisation. (2001). The International Classification of Functioning, Disability and Health, ICF. The International Classification of Functioning, Disability and Health, ICF Retrieved from <http://www.who.int/classifications/icf/en/>

World Health Organisation. (2017, February). 10 facts on the state of global health.

Retrieved from <http://www.who.int/mediacentre/factsheets/fs300/en/>

World Health Organization. (2002). India: - Survey of disabled persons.

<http://www.ilo.org/surveydata/index.php/catalog/196/study-description>

World Health Organization. (2011). Safety and security on the internet. Challenges and advances in Member States Retrieved from

http://www.who.int/goe/publications/goe_security_web.pdf

Yueh, B., Shapiro, N., MacLean, C. H., & Shekelle, P. G. (2003). Screening and management of adult hearing loss in primary care: Scientific review. JAMA, 289(15), 1976-1985.

APPENDIX A

APPENDIX A.1 - ETHICS APPROVAL LETTER



HUMAN ETHICS COMMITTEE

Secretary, Rebecca Robinson
Telephone: +64 03 369 4588, Extn 94588
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2017/19/LR

23 May 2017

Seema Diwan
Communication Disorders
UNIVERSITY OF CANTERBURY

Dear Seema

Thank you for submitting your low risk application to the Human Ethics Committee for the research proposal titled "Evaluation of online hearing-related information in Hindi".

I am pleased to advise that this application has been reviewed and approved.

Please note that this approval is subject to the incorporation of the amendments you have provided in your email of 11th May 2017.

With best wishes for your project.

Yours sincerely

R. Robinson
pp.

Associate Professor Jane Maidment
Chair, Human Ethics Committee

APPENDIX A.2. DEMOGRAPHIC QUESTIONNAIRE

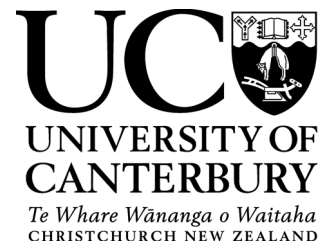
University of Canterbury

Department of Communication Disorders

Private Bag 4800

Christchurch 8140

New Zealand



ID number (*For the researcher to complete*): _____

What is your current age (in years)?

How many years of education have you completed?

Please answer the following questions by circling your choice (or highlighting your choice if you are completing this form electronically).

What is your current gender?

Male

Female

Other

What is your current relationship status?

Single (never married)

Widowed

Married

Divorced

In a committed relationship

Separated

Which ethnic group do you belong to?

New Zealand European

Maori

Samoan

Cook Island Maori

Tongan

Niuean

Chinese

Indian

Other such as Dutch, Japanese or Tokelauan. Please state: _____

APPENDIX A.3. INFORMATION SHEET

Department of Communication Disorders, phone: (03) 369 4519



Email: seema.diwan@pg.canterbury.ac.nz

Evaluation of Online hearing-related information in Hindi

I am Seema Diwan, a 2nd year Master of Audiology student. For my thesis, I am conducting a study evaluating online hearing-related information in Hindi. My thesis will compare how easy or difficult online information about hearing is to read using two methods. First, I will use computer software to calculate how easy or difficult the information is to read. Then, I will ask a group of Hindi speakers to read a sample of the information, as described below.

If you choose to take part in this study, your involvement in this project will be to read 2 samples of online hearing-related information in Hindi. Every 5th word in both samples is replaced with a blank space. Your task is to try to fill in the blanks, using only the information you have from the sample.

There is a risk that you may feel distress when you do this task. You may find this task difficult, because the sample may be difficult to read. If you don't know what word to fill in, that's OK. Just take a guess. You may also feel distress as you read about hearing-related topics. If you feel distress or want more information about hearing loss, you can get help from Healthline: www.health.govt.org or ring 0800 611 116.

Participation is voluntary and you have the right to withdraw at any stage without penalty. You may ask for your raw data to be returned to you or destroyed at any point. If you withdraw, I will remove information relating to you. However, once analysis of raw data starts on 1 August 2017, it will become increasingly difficult to remove the influence of your data on the results.

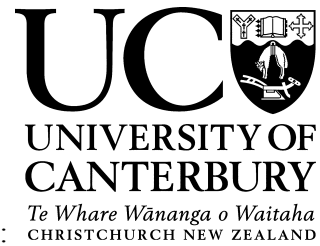
The results of the project may be published, but you may be assured of the complete confidentiality of data gathered in this investigation: your identity will not be made public without your prior consent. To ensure anonymity and confidentiality, I will use a unique ID code instead of your name on all data sheets pertaining to you. I will store your consent form and data sheets in separate, locked filing cabinets in locked and alarmed buildings on the University of Canterbury campus. A thesis is a public document and will be available through the UC Library.

Please indicate to me on the consent form if you would like to receive a copy of the summary of results of the project. The project is being carried out part of my Master of Audiology thesis by me (Seema Diwan) under the supervision of Dr Rebecca Kelly-Campbell who can be contacted at Rebecca.kelly@canterbury.ac.nz. She will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate in the study, you are asked to contact my supervisor by email (Rebecca.kelly@canterbury.ac.nz). She will send you the study materials and you can return them to her by email or via a postage-paid return envelope. She will remove your name from the study material and pass the materials on to me.

APPENDIX A.4. CONSENT FORM



Department of Communication Disorders, phone: (03) 369 4519, email: seema.diwan@pg.canterbury.ac.nz

Evaluation of Online hearing-related information in Hindi

Please read the statements below, and sign if you agree to participate in this research.

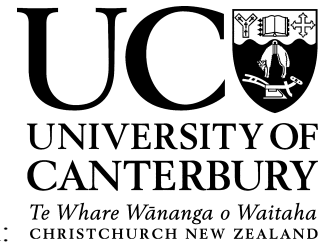
- ☐ I have been given a full explanation of this project and have had the opportunity to ask questions.
- ☐ I understand what is required of me if I agree to take part in the research.
- ☐ I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.
- ☐ I understand that any information or opinions I provide will be kept confidential to the researcher and her supervisor and that any published or reported results will not identify the participants. I understand that a thesis is a public document and will be available through the UC Library.
- ☐ I understand that all data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after 5 years.
- ☐ I understand the risks associated with taking part and how they will be managed.
- ☐ I understand that I can contact the researcher (Seema Diwan: sdi32@uclive.ac.nz) or her supervisor (Rebecca Kelly-Campbell: Rebecca.kelly@canterbury.ac.nz) for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)
- ☐ I would like a summary of the results of the project.
- ☐ By signing below, I agree to participate in this research project.

Name: _____ Signed: _____ Date: _____

Email address *(if you would like a summary of the results)*:

Please return this form and return it to me either by email or in the postage-paid return envelope provided to you.

APPENDIX A.5. CLOZE PROCEDURE



Department of Communication Disorders, phone: (03) 369 4519, email: seema.diwan@pg.canterbury.ac.nz

Evaluation of Online hearing-related information in Hindi

Please read the 2 samples of online hearing-related information below. Every 5th word in both samples is replaced with a blank space. Your task is to try to fill in the blanks, using only the information you have from the sample.

You may find this task difficult, because the sample may be difficult to read. If you don't know what word to fill in, that's OK. Just take a guess.

Sample 1

यदि अवरोध कान के _____ भाग में या सुनने _____ संबंधित नस में है, _____ इसे सेन्सरी न्यूरल डेफनेस _____ हैं। इसके सामान्य लक्षण में _____ से सांय-सांय की _____ अथवा तरह-तरह की _____ आना और कान का _____ होना, कान में दर्द _____, जो मोबाइल फोन के _____ इस्तेमाल से बढ़ सकता _____, चक्कर आना, व्यक्तित्व से _____ मानसिक परेशानियां आदि होते _____। सेन्सरी न्यूरल बहरापन के _____ पैदाइशी बहरापन, जो वंशानुगत _____ पैदा होते समय बच्चे _____ देर से रोने पर _____ में आक्सीजन की कमी _____ कारण अथवा कान के _____ विकसित न होने के _____ हो सकता है। इनके _____ ध्वनि प्रदूषण जैसे तेज _____ के जेनरेटर, प्रेशर हार्न, _____ द्वारा प्रदूषण से भी _____ हो सकता है।

Sample 2

शोधकर्ताओं ने एक शोध _____ ये पाया के जिन _____ के सुनने शक्ति किसी _____ के वजह से चली _____ थी। प्याज़ और पानी _____ मिश्रण सेवन से उनके _____ की शक्ति में

बहुत _____ सुधर हुआ । बहुत सारे _____ जो सुनने की शक्ति _____ चुके थे, उन्होंने अपने _____ में लहसुन लेना शुरू _____ और उनको बहुत ही _____ देने वाले लहसुन के _____ देखने को मिले और _____ में लहसुन खून के _____ को बड़ा कर और _____ के उस हिस्से को _____ आवाज़ को सन्देश में _____ कर शरीर में केमिकल _____ पैदा करता है को _____ ठीक कर के कान _____ सुनने की शक्ति को _____ लाने में मदद करता _____ ।